

Astrophysics Scientific Information Management Policy

Version 1.8

Effective Date: Feb. 14th, 2023

Expiration Date: Jan 13th, 2027

NASA Headquarters
Washington DC

Approved:

Astrophysics Associate Director

Date

Change Log

Change History

<i>Revision</i>	<i>Effective Date</i>	<i>Description of Change</i>

Contents

1. Introduction

2. Purpose and Scope

3. Astrophysics Data Policies

Appendix A. Abbreviations and Acronyms

1. Introduction

The Astrophysics Division of NASA’s Science Mission Directorate (SMD) is dedicated to the study of the universe and everything in it. Astrophysics Division funded space missions, R&A (Research and Analysis), and numerous other investigations power some of the most illuminating explorations of the mysteries that intrigue humanity.

The information produced by these activities results from a significant public investment and the adoption of “Open Science” guided by the FAIR principles ([Findability, Accessibility, Interoperability, and Reusability](#)), aimed at maximizing data findability, accessibility, interoperability, and reusability, is the best way to maximize the return on this investment. Building on the recommendations of the [SMD's Strategy for Data Management and Computing for Groundbreaking Science 2019-2024](#), existing Government directives, NASA policy, community best practices, studies by the national academies, and community-led studies, SMD has created the Science Mission Directorate Scientific Information Policy ([SPD-41a](#)).

The purpose of the Astrophysics Scientific Information Management Policy is to clarify details of the implementation of SPD-41a to the scientific information produced by Astrophysics Division funded activities. This document provides guidance through the complete information lifecycle from measurements and information creation to persistent, accessible curation and sustainment. Scientific information includes, but is not limited to publications, data, and software produced in the pursuit of scientific knowledge .

Data and infrastructure funded by the Astrophysics Division are produced by projects of all scales including missions (spaceflight, etc.), sub-orbital experiments, and research and analysis (R&A) investigations. The proposals for these activities define the science objectives that determine

the required data products. The production and serving of the data from a mission, experiment or R&A activity are governed by an Open Science (this subsumes Data Management Plan (OSDMP). This subsumes the Project Data Management Plan (PDMP) and includes the previously required Data Management Plan (DMP). Beginning with mission Phase E, and continuing after the phase F is concluded, data are stored and served by the Astrophysics Division funded archives, and a duplicate set may also be available through mission or instrument data centers. [Note: in some instances, commissioning data, formally part of Phase D, may also be stored in mission data centers]. Guidance for archiving the data from R&A investigations will be made available in an update to this policy document and will be consistent with the common expectations established for all Federal agencies in “Desirable Characteristics of Data Repositories for Federally Funded Research” (<https://www.whitehouse.gov/wp-content/uploads/2022/05/05-2022-Desirable-Characteristics-of-Data-Repositories.pdf>).

Data from astrophysics missions, sub-orbital experiments, and R&A investigations are to be made available in accordance with the Scientific Information Policy for the Science Mission Directorate (SPD-41a). Implementation of such access can be facilitated by the established NASA astrophysics [data centers](#) which can provide guidance on standardized formatting and hosting. Uniform data access is enabled by the use of standard tools (such as Python) or by APIs that equip researchers to apply user- or community-developed tools for a wide array of research from single instrument to cross-mission analysis and visualization.

2. Purpose and Scope

2.1 Purpose

The purpose of this document is to present the Astrophysics Scientific Information Management Policy. These guidelines are based on the vision and goals of the “[Scientific Information Policy for the Science Mission Directorate](#)” (SMD Policy Document SPD-41a). and those of the Directorate’s [Strategy for Data Management and Computing for Groundbreaking Science](#). However, they capture the particular needs of the community served by NASA’s Astrophysics Division and clarify the applications of SMD data policies to NASA astrophysics activities. The scope of these policies encompasses all phases of a project (including missions, experiments, and investigations) from pre-announcement, through proposal and throughout a project life cycle.

New projects (including missions, experiments, and investigations) shall follow all parts of this policy and SPD-41a. Existing missions, experiments, and investigations should adopt all parts of this policy and SPD-41a that add scientific value and are consistent with available resources. This policy document will be updated on an as-needed basis; major updates will be announced to the astrophysics community through the NASA Astrophysics NSPIRES mailing list (subscribe to the NSPIRES mailing lists by logging in at <https://nspires.nasaprs.com/> and checking the appropriate boxes under "Account Management" and "Email Subscriptions").

NASA’s Astrophysics Division views data users from the general science community as responsible partners in the improvement of the data environment and of the data products themselves.

Two fundamental principles underpin the Astrophysics Scientific Information Management Policy:

1. Implementing NASA's open data policy in a manner driven by scientific efficacy and the minimization of compliance burden. This involves making scientifically useful, high-quality, data publicly available as soon as practical, and
2. Adhering to the goal of early (as soon after acquisition as feasible) and continuing scientific data usability, which requires uniform descriptions of high-quality data products, adequate documentation, sustainable and open data formats, easy electronic access, appropriate analysis tools, and data preservation which aims for 'data immortality'.

In this policy document, all mandatory actions (i.e., requirements) are denoted by statements containing the term "shall." The terms "may" or "can" denote discretionary privilege or permission, "should" denotes a good practice, and is recommended, but not required, "will" denotes expected outcome, and "are" or "is" denote descriptive material.

2.2 Scope

This policy applies to all NASA Astrophysics Division funded activities regardless of the funding vehicle as set out in Section II of the SMD Policy Document SPD-41a.

2.3 Definitions

Consistent with SPD-41a, key terms in this policy are defined as follows:

- **Scientifically Useful:** any material that you as a scientist would need in order to assess, interpret, evaluate, validate, or reproduce the results of a scientific project. Questions and clarifications will be provided by the responsible Program Officer/Program Scientist and the final authority rests with the Division Director. As the Astrophysics Division is gaining practical experience with the application of "scientifically useful" to the broad range

of information circumstances within NASA astrophysics, community input will be sought and community standards will be considered.

- **Information:** Scientific knowledge produced as part of a research activity. This can include, but is not limited to, publications, data, and software.
- **Data:** Scientific or technically relevant information that can be stored digitally and accessed electronically. This includes:
 - Information produced by missions include observations, calibrations, coefficients, documentation, algorithms, and any ancillary information. Further definitions of mission data and examples are provided in Appendix E of SPD-41a.
 - Information needed to validate the scientific conclusions of peer-reviewed publications such as the data required to derive the findings communicated in figures, maps, graphs, and tables.
 - This does **not** include laboratory notebooks, preliminary analyses, intermediate data products, drafts of scientific papers, plans for future research, peer review reports, communications with colleagues, or physical objects, such as laboratory specimens.
- **Publications:** Scientific and technical documents released through print, electronic, or alternative media. This includes peer-reviewed manuscripts, technical reports, conference materials, and books. This does **not** include internal reports, laboratory notebooks, preliminary analyses, drafts of scientific papers or preprints, plans for future research, peer review reports, proposals, or communications with colleagues.
- **Reports:** Documents produced through print, electronic, or alternative media containing Scientific and Technical Information (STI). These documents are usually not peer reviewed. This includes technical publications, technical memoranda, conference publications, special publications, and technical translations. This does not include contractor reports or interim or final research grant reports.
- **Software:** Computer programs, including source and object code, that provide users some degree of utility or service. Further definitions and examples of software are provided in Appendix F of SPD-41a.

- Source code: Human-readable set of statements written in a programming language that together compose software. Programmers write software in source code, often saved as a text file on a computer. The terms code and source code are often used interchangeably.
-

3. Astrophysics Policies

These policies governing the production, distribution, and storage of Astrophysics information were developed in alignment with the data management principles and policies of the Science Mission Directorate.

3.1 The Astrophysics Division is committed to the full and open sharing of scientifically useful data produced by projects of all scales (including missions, sub-orbital experiments, and research and analysis investigations) with all users in an accessible manner as soon as such data become available.

a. A period after mission data have been obtained may be allowed for activities such as initial calibration and validation per the current SMD standard, as defined in SPD-41a, prior to the data reaching the archives. This period will be as short as practicable. For pointed observations this period is typically from days to weeks and for survey observations this can depend on cadence. Clarifications and variances to this include:

(i) Collaborations with international partners will expect minimal exclusive use periods. For future missions involving an International partner, if the Astrophysics Division agrees to extend limited data beyond zero months it will be a rare occurrence, will require justification unique to the mission, will be specified in the MoU (Memorandum of Understanding), and will require a waiver obtained from the SMD Associate Administrator.

(ii) For missions with General Observer/Guest Observer programs, the mission PI/project scientist/observatory director may provide a process for waivers that increase (or decrease) the default data exclusive use period. These waivers should be on a case-by-case basis, infrequent, and always with compelling justification preferably based on panel review evaluation, e.g., to protect the timely completion of a graduate

student's thesis. The HQ program scientist must be informed of all waiver decisions. Waivers that include a cumulative extension to the default exclusive access period greater than 6 months must be submitted to the HQ program scientist for approval. A written record must be maintained by the mission PI/project scientist/observatory director of all waiver decisions, including the terms of and rationale for the waiver decision; although this written record does not have to be publicly posted, it is part of the public record and can be obtained with a FOIA request.

(iii) Mission software should be open-source and must be shared if this is scientifically useful (as defined in section 2.3 above) and practicable. This shall include sufficient documentation to allow reuse. Practicability speaks to situations such as the reuse of well-tested and reliable legacy code (a fairly common practice with space missions) that is written using near-obsolete languages/practices and would require very considerable resources to convert to open source. If rewriting legacy code is impractical, a waiver may be requested using the process outlined in SPD-41a.

(iv) For clarification, we reiterate Section II.B.i.b of SPD-41a that: "This does not include internal reports, laboratory notebooks, preliminary analyses, drafts of scientific papers or preprints, plans for future research, peer review reports, proposals, or communications with colleagues."

(v) For clarification, we reiterate that SPD-41a explicitly excludes export-controlled and proprietary information.

b. Astrophysics research data shall become publicly available no later than the publication of the peer-reviewed article that describes it or at the conclusion of the research award period. Scientifically useful data and advanced data products associated with the award that have not already been made public shall be made publicly available by the conclusion of the research award.

c. Data collected as part of astrophysics crowdsourcing projects or citizen science projects shall be made public.

3.2 Astrophysics Division funded archives will plan and follow data acquisition policies that ensure the collection and usability of long-term mission data sets needed to satisfy the research requirements of the Astrophysics community. For smaller research projects, the archives will coordinate with SMD core services to secure the availability of resources to host the scientifically useful research products.

a. Missions funded by the Astrophysics Division shall make available any and all scientifically useful data products along with any calibrations, coefficients, algorithms, documentation, and ancillary data used to generate these products.

i. All data that is in a state that allows a complete reanalysis of the raw data, shall be archived along with tools (including instructions on their proper usage) for its reduction to science products and documented algorithms for this process. Relevant engineering and “housekeeping” data should also be preserved. However, each of these requirements must meet the standard of “scientific usefulness” to justify the effort and resources necessary for archiving.

b. The Principle Investigator of research investigations shall make available any and all scientifically useful research data products along with algorithms, documentation, and ancillary data used to generate these products.

c. Astrophysics archival data shall be made available in convenient, modifiable, and open formats. The data and metadata shall follow as much as possible the specifications for formats and contents provided by the data center or journal that will archive and serve them. Best practices are described in <https://arxiv.org/pdf/2106.01477.pdf>.

d. Scientifically useful data from Astrophysics Division funded activities shall follow the FAIR Guiding Principles for scientific data management and stewardship. This means data should be findable, accessible, interoperable, and reusable (FAIR; Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3:160018 doi: 10.1038/sdata.2016.18 (2016).

e. Data providers (missions, archives, and PIs that publish primary data products) shall register their datasets with Digital Object Identifiers(DOIs) to facilitate their citation and access. They will work with ADS (Astrophysics Data System; <https://ui.adsabs.harvard.edu/>) to ensure that these data products are properly represented in its database, and they will provide guidance to their users on how to link and cite them in the literature. Scientists re-using data from missions and archives will use the appropriate DOIs to properly acknowledge their use, either via a mention or a citation in their peer review manuscript. Authors are also encouraged to do so for e.g. preprints, conference proceedings, and books.

f. Astrophysics data shall be reusable with a clear, open, and accessible data license (e.g., Creative Commons license). If there are no other restrictions on the data, the data shall be released with a Creative Commons Zero license.

h. Software funded by the Astrophysics Division shall be released under a permissive license that has broad acceptance in the community and shall be made available in a publicly accessible repository that is widely recognized by the community. If commercial software is used, then the Astrophysics Division-funded executable code shall be released and made available.

i. All Astrophysics Division funded publications shall be made publicly accessible. An openly accessible version of the as-accepted, peer-reviewed manuscripts shall be available via a NASA-designated repository at the time of publication. Publishing the manuscript as

Open Access and posting a version on a community recognized preprint server are encouraged. We note that most publications used by our community have already signed up to CHORUS (<https://www.chorusaccess.org/>) making compliance essentially automatic. The cost of making publications publicly accessible is an allowable research expense; the cost of complying with SPD-41a and this policy document should be included in the budget request of proposed investigations.

3.3 The Science Mission Directorate (SMD) will collect a variety of metrics intended to measure or assess the efficacy of its data systems and services and assess user satisfaction. Consistent with applicable laws, SMD will make those data available for review. See Section X of SPD-41a.

3.4 Archives funded by the Astrophysics Division will continue a principle of non-discriminatory data access to create an equitable playing field for all users. For data products supplied from an international partner or another agency, the Division will restrict access only to the extent required by the appropriate agreements.

3.5 All NASA astrophysics missions, projects, and grants and cooperative agreements shall document their implementation of these data management policies.

- a. Missions shall generate a Project Data Management Plan (PDMP) that describes how information will be generated, processed, distributed, analyzed, and archived.
- b. PI led missions shall create a Data Management Plan (in Explorers AO).
- c. Research projects shall generate an Open Science and Data Management Plan (OSDMP) that will include a Data Management Plan, Software Management Plan, and other Open Science activities.
- d. For extended missions, compliance with this policy will be part of their Senior Review. For PI-led prime missions, compliance shall be included with their quarterly/annual reports.

3.6 For the code of conduct and guidelines on how to make contributions to astrophysics funded software project (see section III.D.iv of SPD-41a) please refer to the [Astrophysics Division “Statement of Principles”](#)

3.7. We reiterate that the SPD-41a “Appendix F. Guidelines for sharing software,” for PI-led software (single use, libraries, analysis, model and simulation software), guidelines apply to scientifically useful products as defined in 2.3. There is a variance process that will take into account difficulties faced by e.g. students/early career scientists.

3.8. Variances: Guidance on obtaining variances on policies for Research Information is provided in section VII.E of SPD-41a. The process for obtaining variances in general are provided in section IX of SPD-41a. Also see Section 3.1.a.ii above.

Appendix A. Abbreviations and Acronyms

AA	Associate Administrator
AAS	American Astronomical Society
ADS	Astrophysics Data System
AGU	American Geophysical Union
AO	Announcement of Opportunity
APS	American Physical Society
BSD	Berkeley Software Distribution
APD	Astrophysics Division
CAN	Cooperative Agreement Notice
CDF	Common Data Format (CDF)
DMP	Data Management Plan
DOI	Digital Object Identifier
DPI	Digital Persistent Identifier, synonymous with PID
EAR	Export Asset Regulations
FAIR	Findable, Accessible, Interoperable, and Reusable
FAR	Federal Acquisition Rules
FITS	Flexible Image Transport System
HEASARC	High Energy Astrophysics Science Archive Research Center
HIPAA	Health Insurance Portability and Accountability Act
IDL	Interactive Data Language
IPAC	Infrared Processing & Analysis Center
IRSA	NASA/IPAC InfraRed Science Archive
ITAR	International Traffic in Arms Regulation
MAST	Barbara A. Mikulski Archive for Space Telescopes
MoU	Memorandum of Understanding
NED	NASA Extragalactic Database
NetCDF	Network Common Data Format
NExScI	NASA Exoplanet Science Institute
NPD	NASA Policy Directive
NPR	NASA Procedural Requirements
NRA	NASA Research Announcement

NSTC	National Science and Technology Council
OMB	Office of Management and Budget
OPEN	Open, Public, Electronic, and Necessary
ORCHID	Open Researcher and Contributor ID
OSS	Open Source Software
OSTP	Office of Science and Technology Policy
PDMP	Project Data Management Plan
PID	Persistent Identifier, synonymous with DPI
R&A	Research and Analysis
RFC	Request for Comment
RFI	Request for Information
ROSES	Research Opportunities in Space and Earth Science
SALMON	Stand-Alone Missions of Opportunity Notice
STScI	Space Telescope Science Institute
SMD	Science Mission Directorate
SOS	Subcommittee on Open Science
SPD	SMD Policy Document