Planetary Data System
Mission Proposer's Archiving Guide (MPAG)

September 21, 2016

Version 4.0.0

Contents

CHANGE LOG

Revision	Date	Description	Author
Start Draft	April 2, 2003	Initial outline	R. Joyner
Version 2.0	2014-05-1	Updated to comply with PDS4	R.Joyner

Version 2.10	Date	Description	Author
Section 8	2015-10-19	Removed costing section pending rewrite that is valid for PDS4 (MC AI 2015-09-28/01)	Simpson
1.2 6.2 Appendix B	2015-10-19	Removed references to costing section and web page with costing procedure	Simpson
1.2	2015-10-19	Corrected reference to Appendix C (now to B)	Simpson
1.4	2015-10-19	Updated list of applicable documents	Simpson
2.0 3.5.2 3.5.4 6.5.1 7.0 7.5 Appendix A	2015-10-19	Minor grammatical, wording, font, and pagination corrections and improvements	Simpson
4.1.2	2015-10-19	Noted upcoming node name changes at Rings and Imaging	Simpson
7.5	2015-10-19	Added sentence noting that EN is the steward for all context products.	Simpson
7.7	2015-10-19	Removed section heading (there is no content)	Simpson
Appendix B	2015-10-19	Updated references 3 and 4, removed 5 and 7; renumbered remaining references.	Simpson
4.1	2015-11-27	Changed “Management Office” to “Project Office” and updated names of CIS and RMS nodes to reflect nomenclature as of December 2015. Corrected sentence structure by adding verbs to DN summaries.	Simpson

Version 4.0	Date	Description	Author
Rewrite	2016-09-21	Restructured and updated to be mission oriented.	T. King

1 Introduction

NASA's Planetary Science Division supports a wide diversity of projects ranging from investigations by individual researchers to large flagship missions. Across the full spectrum of projects NASA requires the data and other resources that are produced to be preserved in the Planetary Data System (PDS) or an equivalent archive. This guide describes the process and requirements when archiving data with the PDS. The basic PDS archiving requirements are the same for all types of projects. However the roles, detailed steps, and responsibilities will vary from one project to another, typically related to the scale of the project. This guide is aimed at NASA's competed flight missions[1]. Being familiar with PDS concepts and the general design of a PDS compliant archive is key to writing a successful Data Management Plan (DMP)[2] in a mission proposal. This guide describes the items that you, as a mission proposer, need to consider while preparing a proposal to a NASA flagship, explorer or discovery class program. This guide also provides a brief summary of information found in PDS standards and will help you to plan and estimate the effort of preparing your archive for submission to the PDS.

2 PDS Background

The PDS archives and distributes scientific data from many sources, including planetary missions, astronomical observations, and laboratory measurements. Its purpose is to ensure the long-term access and usability of data and to support and stimulate advanced research. All PDS data are peer-reviewed, well-documented and publicly available. That is, the data may be exported outside of the United States under the "Technology and Software Publicly Available" (TSPA) classification.

The PDS is a federation of teams with expertise in different science disciplines. Each team is called a "node." Within the PDS, there are nodes focusing on the scientific disciplines of Atmospheres, Geosciences, Cartography and Imaging Sciences, Planetary Plasma Interactions, Ring-Moon Systems, and Small Bodies. Additionally, there are two support nodes of the PDS: the Engineering Node and the Navigation and Ancillary Information Facility. For a current list of contacts for each node see

https://pds.nasa.gov/contact/contact.shtml

3 A Typical PDS Archive

A PDS archive typically includes data, labels containing information about the content and format of the data, and documentation about the data and mission. Together the labels and the documentation must be sufficient to enable individuals to find, read and use the data. The documentation should include descriptions of which instrument was used to make the observation, what platform was the host of the instrument and which objects were observed. Collectively these provide the context of the observation. Additional documentation may include papers or other documents that describe the instrument, calibration steps or processing that was performed. For the data, the structure and format must be described in sufficient detail to enable others to read the data. Both the data labels and the context documentation are specified with PDS-defined metadata in compliance with PDS standards. In particular, the standards require that data products conform to the structures and formats that can be

[1] There is a separate guide for R&A projects called the Individual Proposer's Archive Guide (I-PAG)
[2] A DMP may be known by other names such as Data Management and Archive Plan (DMAP).

described by PDS-defined metadata. Data products that can be read only by external software (public or proprietary) are not compliant. To meet archiving standards data must be readable using only the required metadata in the PDS label. The only exception to this rule are SPICE kernels, as defined by NASA's Navigation and Ancillary Information Facility (NAIF) node. The context documentation, other documentation, labels and data are logically grouped into an archive, which is placed in the PDS, made available to users, and maintained indefinitely. Every archive submitted to the PDS must pass a peer review prior to its public release.

4 PDS4 Details and Definitions of Terms

PDS4 specifies the standards and services used by PDS. PDS4 is an integrated system designed to improve discoverability and access to data across multiple storage locations. Each item stored in a PDS archive – whether data, documentation, or other material – is considered a product. A product consists of the archived item and the label containing its metadata. Every product is assigned a unique identifier so it can be referenced and managed. A product is usually a single file, but it is possible for a product to consist of multiple files described by a single label.

Products are organized into Collections and Collections are organized into Bundles. While designing a PDS4 compliant archive it is important to keep in mind the following general principles.

Everything is a product.

A product is any item in the archive along with its label. Each product is assigned an identifier that is unique throughout all of PDS. Products may be data, documents, context files, XML schema or other digital objects.

Collections are groups of similar products.

A collection consists of products that share one or more attributes. These attributes can be the type of product, a time range, acquisition method (i.e., instrument), processing level or any other logical attribute. Every archived product must belong to a collection, so every archive submission to PDS will have at least one collection. A product may belong to more than one collection, in which case it is called a secondary member of the additional collections.

Some examples of collections are:

- Document Collection – Contains documentation necessary for understanding and using the data
- Data Collection – Contains observed, derived or generated data, logically grouped by shared attributes
- Calibration Collection – Contains calibration data separated from the data collection(s) and/or more detailed information about complex calibration efforts.

Bundles consist of related collections.

A bundle contains collections that are related by a high level concept such as project, acquisition method, processing level, or utility. The collections in a bundle are complementary to each other. Every archive submission to PDS will have at least one bundle.

For example, a bundle may include:

- all collections (data, document, calibration) for a particular instrument, or
- data collections from a single lab or lab project, or
- all collections from a single field campaign, or
- all collections from a research or analysis project.

The relationship among products, collections and bundles is shown in Figure 1. In a PDS4 archive a bundle contains PDS4 metadata stored as XML documents which describe the collections that are part of the bundle. Each collection contains XML documents and related files which describe the products that are included in the collection. Each product consists of an XML document and related files.

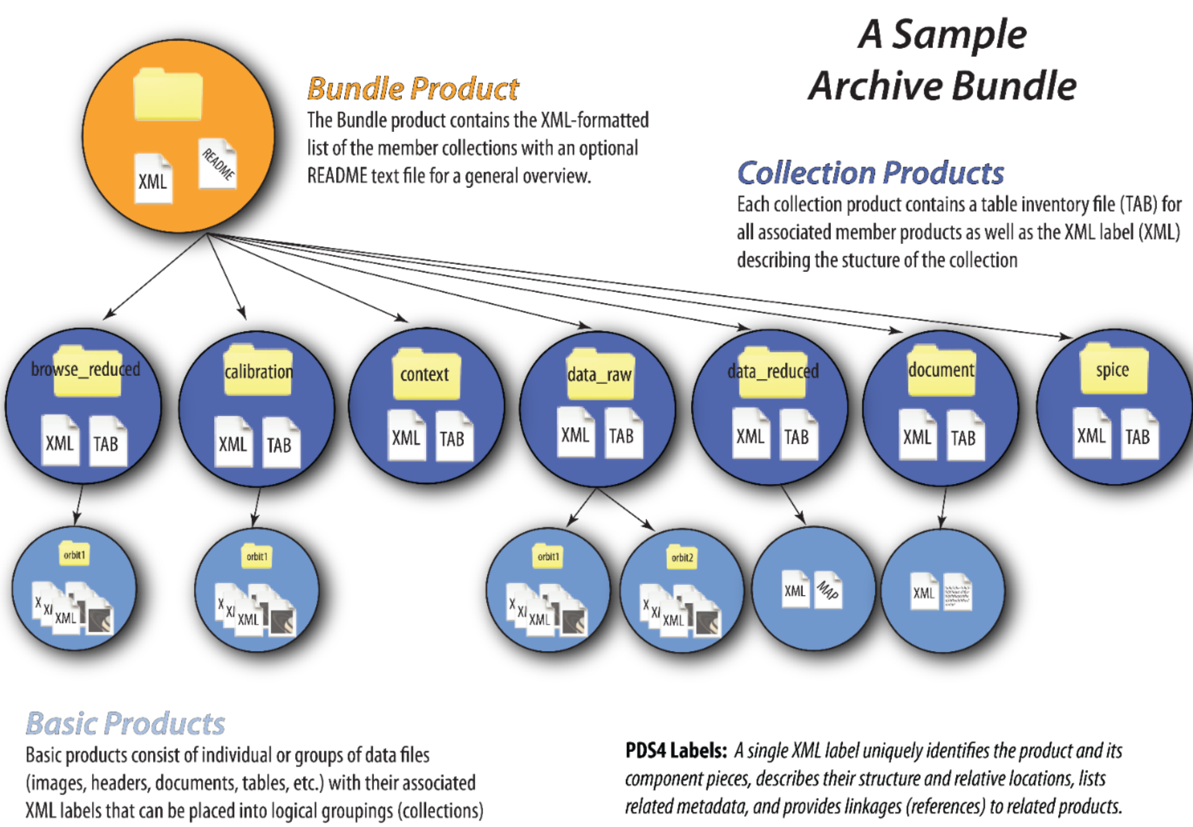


Figure 1 Relationship of product, collection and bundle.

5 PDS Archiving Process

The PDS archiving process consists of multiple steps, beginning with identifying the data products, designing the collections and bundles, and continuing with the creation of context products, labeling of products, delivery of the products to PDS, conducting a peer review and resolution of any issues discovered during peer review. If an archive is incrementally delivered to the PDS there may one or more delta reviews to ensure the entire archive collection meets PDS standards and requirements.

During or prior to the design phase, mission and instrument personnel should define the data products they intend to archive, estimate their volume and generation rate, and negotiate a preliminary delivery procedure with the PDS. Adequate documentation for both understanding and using the data is critical to each archive; mission planners should be mindful of the archive documentation requirement as they develop explanatory materials for other purposes. A suitable archive includes complete descriptions of the operation of each instrument and of the calibration procedures that are suitable for general release (ITAR/EAR cleared). It will include raw and processed data along with all files needed to carry out the calibration. The mission proposal call and/or NASA managers may require specific special or higher level products. These might include derived products such as mosaics, maps, data resampled to a common interval, fits to model results, or inversions. These products will also need to be included in the archive. Additionally, a mission may determine that there are other products (combinations of data from multiple sources, the results of analyses not originally planned, etc.) that are worthy of archiving. The PDS encourages inclusion of these special products in the archive.

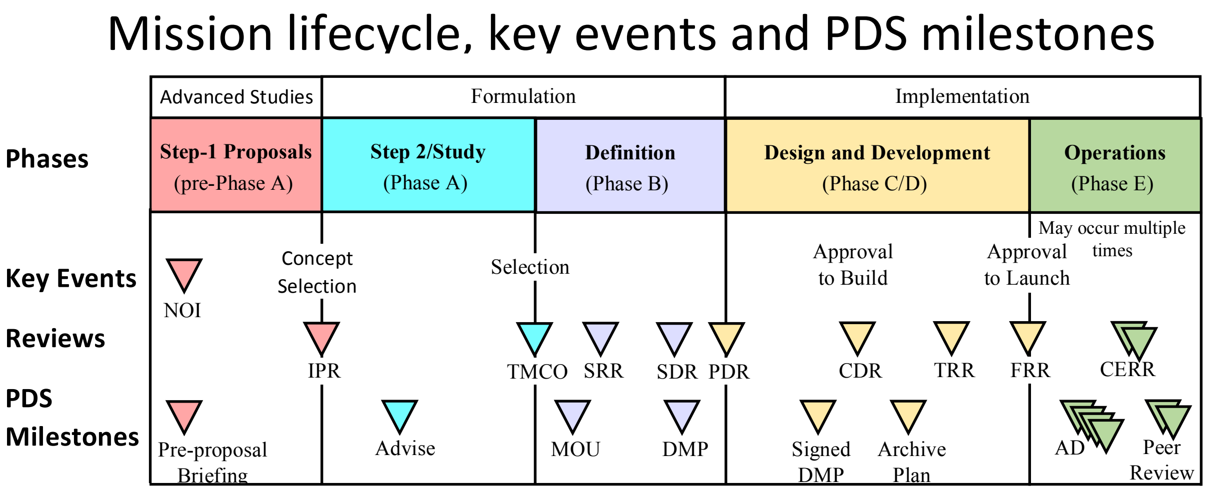
For products that rely on cartographic coordinate system - for example maps, mosaics and other geospatial products - the coordinate system must follow standards and practices of the Working Group on Cartographic Coordinates and Rotational Elements (WGCCRE) of the International Astronomical Union (IAU). These standards and practices are spelled out in online statements (https://pds.nasa.gov/documents/sr/Chapter02.pdf; and http://sbn.pds.nasa.gov/data_sb/resources/coordinate_systems.shtml). In general, in order for a cartographic coordinate system to be used in a product the coordinate system must be described in detail in publicly released and perpetually available documents. For example, as a white paper archived in the PDS or as a paper in a scientific publication. Such a coordinate system must be consistent with that of data already in the PDS for the target body (or bodies), must have a prime meridian which is defined by using a clearly identifiable physical feature, and must be consistent with any established IAU coordinate system for the same body.

In addition to acquired data, the mission is expected to archive complete geometric details from launch through end of mission. These typically include the full ephemeris of the spacecraft and orientation of the spacecraft and all instruments, the relationship of these to each other and to coordinate systems on the target and a history of all significant spacecraft events. Additionally, housekeeping data (such as temperatures and power levels) that might be useful in understanding the behavior of instruments may be archived. The mission should also archive ancillary data that are important to either mission planning or interpretation of the data from the mission. These might include contemporaneous, Earth-based observations or key models, such as shape models, used in interpreting the data. Furthermore, in the case of radio science experiments, radiometric tracking data are also archived.

6 PDS Involvement in Mission Planning and Operations

The PDS recognizes that it takes some effort to learn how to produce an archive that conforms to PDS4 standards, and we will help you with this. The task of creating an archive is shared between the mission and the PDS. Through all phases of a mission PDS is a resource you can rely on for advice and guidance. This includes all the steps in archiving the data from preparing a proposal, to creating a Data Management Plan (DMP), to the operation of the mission and making deliveries to the archive. At every step, from preparing a proposal to mission operations, to archive preparation feel free to communicate with PDS as often as needed to ensure that you approach archiving in an efficient manner. We can assist you in identifying the components for your planned archive, assist in developing templates for labels, and help you organize the collection and bundle structure. While writing a proposal feel free to contact the PDS Node most closely aligned with the primary goals of the project. After selection, the PDS Management Council will assign a discipline node as the lead node. In many cases this will be the same node you interacted with during the preparation of the proposal.

In the following sections we describe what PDS will provide at each milestone [1][2] over the course of a project. While the terminology used in this document may vary somewhat from one program to another, missions typically have periods of advance studies, formulation and implementation. Figure 2 shows the various phases for a mission and the PDS-related milestones. Some projects may have additional milestones or documents which are produced and may produce multiple documents to meet a milestone. In some cases, milestones involve PDS signing a document, such as the archive plan sections in a project's DMP, or certification by PDS, such as in a peer review.



NOI – Notice of Intent
IPR – Initial Proposal Review
TMCO - technical management, cost, and other
SRR- System Requirements Review
SDR - System Design Review
DMP – Data Management Plan
MOU – Memorandum of Understanding

PDR – Preliminary Design Review
CDR –Critical Design Review
TRR – Test Readiness Review
FRR – Flight Readiness Review
CERR – Critical Event Readiness Review
AD – Archive Delivery

Figure 2: Mission lifecycle, key events and PDS milestones.

6.1 Pre-proposal Briefing

Prior to proposal deadlines for NRAs and AOs, NASA often organizes a pre-proposal briefing to provide guidelines and answer questions. As part of this briefing, PDS will normally make a presentation on how PDS works and what an archive requires. The *Mission Proposer's Archiving Guide* (this document) will be provided at the briefing.

6.2 Proposal Data Management Plan (DMP)

A mission proposal is usually required to include a Data Management Plan (DMP). While the specific instructions for the DMP may vary for each mission Announcement of Opportunity (AO), typically a DMP prepared for a proposal describes how the data will be produced, how and by whom it will be managed and how it will be archived. After a mission is selected a more detailed DMP will be created which will be mutually agreed to by the project and the PDS with final approval by NASA (Project Management office and the Program Scientist).

Be sure to check for any requirements specific to the program to which you are proposing because some programs may have additional or more specific requirements for the contents of the DMP.

The PDS recommends the following topics for your DMP:

- Identify the PDS node that will lead the PDS side of the data archiving and any other PDS nodes that will interface with instrument or component teams. Include a letter of support from each of the node(s) stating that you have discussed your proposed data products and that the node(s) will work with you to archive them.
- Discuss the roles and responsibilities of team members in accomplishing the archiving goals.
- Describe a schedule for design, product generation, validation and delivery of your products. Be sure to include time required for participating in all steps of the archive process (see Section 7).
- Include a general schedule for submitting the mission archives for review and release to the science community. A mission of long duration should plan to release archives at regular intervals over the course of the mission. Unless the AO specifies otherwise, the release schedule should allow for the accumulation over no more than three months of each batch of data to be released, and the release should occur, at the latest, by the time the newest data in the batch are no more than six months old.
- Outline the design of your archive, including the bundle(s) and collection(s) as well as any derivatives of the delivered products (i.e., browse images), observational geometry information or supplemental products
- Summarize the products you intend to archive, including the data products, documentation, geometry and ancillary information (e.g., information on how the data were obtained, processed, calibrated---include anything that a user would need to know to use your data as a scientific product). Each type of product should be named and described in terms of processing level, format, units of measure, and estimated volume. Table 1 shows the PDS processing level names and definitions. It is not necessary to submit products at every processing level, but the raw data and at least one level of calibrated data are expected. Necessary documentation is in addition to the data products.
- Discuss how geometry information will be generated. For example, whether SPICE will be used or by some other method.
- Provide estimated data volume. The estimate should include accumulations rates (bytes by year) and be sub-divided by instrument.

Processing level	Description
Telemetry	An encoded byte stream used to transfer data from one or more instruments to temporary storage where the raw instrument data will be extracted. PDS does not archive telemetry.
Raw	Original data from an instrument. If compression, reformatting, packetization, or other translation has been applied to facilitate data transmission or storage, those processes are removed so that the data values can be readily used. These are often called Experimental Data Records (EDRs).
Partially Processed	Data that have been processed beyond the raw stage but that have not yet reached calibrated status. These and more highly processed products are often called Reduced Data Records (RDRs).
Calibrated	Data converted to physical units, which makes values independent of the instrument.
Derived	Results that have been distilled from one or more calibrated data products (for example, maps, gravity or magnetic fields, or ring particle size distributions). Supplementary data, such as calibration tables or tables of viewing geometry, used to interpret observational data should also be classified as 'derived' data if not easily matched to one of the other three categories.

Table 1. Data Processing Levels

6.3 Advise and Assist

Early in the mission life cycle, PDS will schedule an orientation for mission representatives to define roles and responsibilities and to establish these points of contact:

- Mission contacts
- Instrument / Science Team contacts
- PDS Discipline Node contacts

A new mission is usually assigned to a single PDS discipline node, which serves as the lead node for the remainder of the mission. Lead nodes, in turn, arrange for support of individual instrument teams through other discipline nodes as appropriate. This includes support for potential data contributors without specific instruments, such as Interdisciplinary Scientists. The lead node works with the mission to ensure that archive plans are developed, the archive designs meet standards and facilitate cross-instrument and cross-mission comparisons, and archive production and validation run smoothly. Many of these elements will be in the detailed DMP. Typically development of the detailed plans is done via a Data Archive Working Group (DAWG) led jointly by the PDS lead node and the mission archiving lead. DAWG membership includes archive representatives from each instrument team and each PDS node that will receive data from the mission.

6.4 Memorandum of Understanding (MOU)

Early in Phase B, the mission and PDS will draft a Memorandum of Understanding (MOU) that defines and delineates the roles and mutual responsibilities regarding archiving and distribution of mission data. For small missions NASA management may waive the need for an MOU. In the MOU the PDS Management Council will assign a discipline node as the lead node.

6.5 Data Management Plan (DMP) and Software Interface Specifications (SIS)

Once a mission enters the development phase it will be required to generate a detailed Data Management Plan (sometimes called an Archive Plan or similar title) as a mission deliverable. A draft of this plan is usually required for the mission Preliminary Design Review. The proposal's Data Management Plan will be a useful starting point for the mission Data Management Plan, but its specific contents will depend on mission requirements.

The lead node will work with the mission to complete the portion of the DMP related to PDS archives. The DMP must contain sufficient detail about the data flow, archival products, and procedures for generating the archives; must clearly identify the individual responsibilities; and must lay out the schedule for archive deliveries and release to the science community. Additionally, the DMP must provide sufficient detail to enable both the mission and the PDS to develop reliable estimates of the personnel and other costs for their respective activities.

Each instrument team on the mission will be required to document the data products to be generated by their instrument in one or more Software Interface Specification (SIS). These documents include enough detail about the data processing, product structure, format, and contents to enable software to be developed to read and understand the data. The SIS documents help satisfy the PDS requirement for documentation and therefore are included in the archive. Each team will work with its designated PDS node to write one or more SIS documents to describe its data products.

Observation geometry data are essential for properly interpreting mission data and must be included in the archive. A program AO may or may not specify the means for generating geometry data, but if you are free to choose, both PDS and the International Planetary Data Alliance (IPDA) recommend that a mission use NASA's SPICE system to produce observation geometry parameters supplied with science data products. If a mission does produce SPICE data, a comprehensive set of mission SPICE kernels must be archived at the NAIF node of the PDS. If a mission produces some other form of geometry data, that geometry data along with adequate documentation must be included in the science data archive.

Note: Some projects may combine the SIS information into the DMP and thereby have a single document. However, the DMP is a signature controlled document, while the SIS is not, so in all but the simplest cases the two documents should not be combined.

6.6 Archive Delivery

Nearly all missions collect data over long periods of time. However, NASA policy considers all scientific data returned from NASA missions to be immediately in the public domain. NASA also recognizes that in order to produce scientifically useful data products there needs to be a period of exclusive access by the mission teams. Barring exceptional circumstances, the exclusive access period may not exceed six months after receipt on the ground of all data within a batch of data. A block of data may contain up to three months of data from the time of downlink of the earliest data in the block. This means that the nominal schedule for delivering data to an archive is every three months (quarterly), with each delivery containing three months of data, and delivered six months after the most recently acquired data in the block. Each delivery must occur long enough before the release date to allow PDS to validate the delivery, typically two or three weeks. The schedule for deliveries will be detailed in the DMP. The deliveries may be made through a science operations center (SOC) or the individual instrument teams.

Each delivery will contain products and other information detailed in the DMP and in conformance with the SIS. Upon receipt PDS will acknowledge the delivery and check the data for compliance.

6.7 Peer Review

PDS requires that all data submissions from data providers pass a peer review before being archived. The PDS coordinates all peer reviews. During a peer review, reviewers will evaluate the usability of the data by utilizing the metadata (in labels) to read and perform a scientific assessment of the data. In addition, the completeness and accuracy of other information in the archive submission will be assessed. Any issues discovered during a review are expressed formally as a “lien” against the products. Major liens are those severe enough to render a product to be unsuitable for scientific research. Minor liens are those which may affect ease of use, completeness of documentation, quality of information (i.e., typos in documentation or metadata) or other similar issues. All liens must be addressed before data products can be archived. A data provider is required to see the process through to having the products archived, including allotting enough time for iterative work on the preparation of products, supporting a peer review and completing any lien resolution. After all liens are resolved there will be a delta review to confirm that the liens were addressed. The PDS will then show the data as "certified" to indicate that the data was peer reviewed and deemed to meet PDS archiving standards.

For short lived missions where data is delivery one or two times, the peer review will occur after each data delivery. For missions that operate over a long period of time and (typically) produce large amounts of data that are delivered incrementally over the course of the mission to the PDS it is preferable to have the data processing pipeline "certified" by PDS. For PDS to be able to certify the pipeline there must be a well-defined set of processing steps with version control on software components used in the data production, and delivered data must indicate which version of the pipeline was used to produce the data. To certify a pipeline PDS will review the product design documentation and sample products before any deliveries are planned to be done. As the project's data processing pipeline matures, and prior to first archive delivery, the PDS will review the products that come out of the pipeline. The products are compared to the peer-review-approved product design documentation to ensure that everything is as expected. In addition, the recipient PDS node will validate the delivery for PDS compliance. If it is demonstrated that the pipeline reliably produces compliant products then a full peer review is not conducted with every delivery. This greatly improves the efficiency of the archiving process. If changes are made to pipeline, whether by changes in software or configuration, the pipeline will need to be re-certified before data can be routinely archived.

7 Estimating Effort and Cost of Creating a PDS Archive

The task of creating an archive is shared between the mission and the PDS.

To coordinate a mission archive, a mission will typically designate an individual to serve as the Mission Archive Scientist. This individual oversees the archiving activities of the entire mission. The Mission Archive Scientist is the primary contact between the mission and the PDS lead node.

Potential tasks and responsibilities for the Mission Archive Scientist include:

- Prepare the archive portions of the Data Management Plan.
- Prepare the descriptive text about the mission, instrument host (e.g., spacecraft), and references to the context products for the PDS registry.

- Ensure the timely compilation by mission team members of context products specific to individual instruments (e.g., instrument, spacecraft, and personnel files).
- Ensure the timely compilation by mission team members of other Archive Description Documentation components, as needed.
- Ensure that any copyright issues have been addressed if previously published material is to be included in the archive.
- Coordinate review of documents, calibration information, and other parts of the archive which may be sensitive for ITAR (International Traffics in Arms Regulations) compliance.
- Ensure that all files delivered to the archive are ITAR cleared.
- Ensure compatibility of formats and definitions from one instrument to another.
- Coordinate delivery schedules for the archival products from different instruments, negotiating and documenting schedule changes as needed.
- Ensure that all data preparers are meeting their obligations/milestones.
- Coordinate the mission team reviews of the archiving plans and details to ensure that all relevant information is captured in the archive.
- Coordinate appropriate mission team participation in the PDS organized peer reviews.
- Organize the mission responses to any liens identified by the review processes.
- Arrange any desired technical workshops for the mission.

Additionally, the mission will identify one or more individuals who will prepare archive deliveries for each instrument. For some projects this activity is coordinated by the Mission Archive Scientist. The preparation of archival products requires a mix of scientific expertise and expertise with data formats and standards. Sometimes a single individual can fill both roles. Archive preparers will have early and continued interaction with PDS personnel because archive deliveries will occur throughout the life span of the mission.

During operations, the general division of effort is as follows:

The mission performs the following tasks (using mission funds):

- Produces data products in acceptable PDS format.
- Produces PDS labels.
- Writes supporting documentation.
- Organizes data, labels, documentation, etc. into collections and bundles.
- Validates labels using PDS provided tools.
- Participates in peer review (may be done either face-to-face or via web, email and phone).
- Makes updates to the products, as necessary, based on peer review recommendations.
- Participates in delta review (if needed).
- Delivers archive packages to PDS on a pre-determined schedule.

The PDS Node performs the following tasks (using PDS funds):

- Provides advice on PDS standards and requirements.
- Assists in designing PDS labels, if needed.
- Helps create context products.
- Provides available PDS tools and guidance on use.

- Accepts the packages and tracks deliveries to PDS.
- Sets up and conducts a peer review.
- Integrates the data into its archives, including making it available on the Node web site.

We have found that this division of tasks helps to have an efficient archiving process; however, there may be situations where this division of tasks is not practical or fully achievable within the scope of the project. PDS is here to help, so contact the appropriate node and discuss how they might assist you.

Our experience indicates that preparation of a simple archive, consisting of one bundle with a few collections with a limited number of products and modest data volume, may take one person up to a month of effort spread over the full duration of the project. For a mission each instrument will have one or more bundles and each bundle will have one or more collections. Additionally, a typical mission delivers data to the archive over an extended period of time. Each archive delivery takes time to assemble, deliver and review. It can take several (2-3) weeks to prepare and complete a delivery to PDS. To calculate the amount of time needed to create the archive one needs to consider how many unique collections will be created. If creating multiple similar collections it will take some amount of time for the first collection and a fraction of that time for each new collection (incrementally more time per collection than the first collection), whereas creating collections that are each unique will take about the same amount of time to create each collection (proportionally more time).

During the course of a mission the time spent on creating a PDS archive is spread across the design, product generation, validation, product delivery, peer review, lien resolution and delta review. In general, more time is spent on product generation and validation than the other phases. Lien resolution can be time consuming if the product generation goes astray, but PDS4 is designed to minimize this likelihood and the PDS Nodes will give you guidance on how to avoid any surprises. PDS also provides tools to aid in the design, generation and validation of products (see https://pds.nasa.gov/pds4/software/index.shtml).

PDS is your partner in preparing an archive. While preparing an archive you should:

- Expect to communicate often and to iterate with a PDS Node on archive design and product formats, labels, etc.
- Expect to validate your data submission using PDS-supplied software.
- Expect to participate in product design reviews and sample data product reviews.
- Be prepared to participate in a peer review of your data submission, with the support of PDS Node personnel.
- Be prepared to resolve any liens.
- Be prepared to participate in a delta review of your data submission.

To illustrate how to estimate the archiving efforts consider this example. A mission has five instruments with each instrument delivering two bundles (raw and calibrated). Each bundle will contain one collection that accumulates over the course of the mission. The mission will take three years to complete resulting in 12 quarterly deliveries to PDS. The minimum amount of time a mission should allocate for designing an archive and preparing for the first delivery is 10 months (1 month for each collection for 5 instruments). An additional two months (2 weeks per quarter) will be needed each year of the mission to package, delivery and review the quarterly deliveries to the PDS. For this example, a total of 16 months, distributed over the duration of the mission, will be needed to archive the data. This was a simple example and many mission archives are more complicated, requiring more time to be allocated for archiving.

8 Additional Information

We understand that not all data providers have extensive backgrounds in PDS standards and processes. While this guide reflects the information found in a variety of PDS documents and specifications, you should refer to these additional resources for more details and suggestions:

- PDS4 Concepts (https://pds.nasa.gov/pds4/doc/concepts)
- PDS4 Standards Reference (https://pds.nasa.gov/pds4/doc/sr/current)
- PDS4 Information Model Specification (https://pds.nasa.gov/pds4/doc/im/current)
- Small Bodies Node PDS4 Wiki (http://sbndev.astro.umd.edu/wiki/SBN_PDS4_Wiki)

9 Appendix A: Glossary

archive: A place in which public records or historical documents are preserved; also the material preserved — often used in plural. Sometimes capitalized when referring to all of PDS holdings — the PDS Archive.

bundle: A set of collections. For example, a bundle could include a collection of raw data obtained by an instrument during its mission lifetime, a collection of the calibration products associated with the instrument, and a collection of all documentation relevant to the first two collections.

collection: A set of products, all of which are closely related in some way.

data object: A physical, conceptual, or digital object.

data provider: A person or organization that assembles archival data for delivery to PDS.

identifier: A unique character string by which a product, object, or other entity may be identified and located. Identifiers can be global, in which case they are unique across all of PDS (and its federation partners). A local identifier must be unique within a label.

information model: A representation of concepts, relationships, constraints, rules, and operations to specify data semantics for a chosen domain of discourse.

label: The aggregation of one or more description objects such that the aggregation describes a single PDS product. In the PDS4 implementation, labels are constructed using XML.

label template: A text file which serves as a pattern for constructing labels.

lead node: One of several consulting nodes designated as the PDS coordinator and primary contact with a project.

lien: A formally expressed issue with one or more products in a collection. Minor liens are those intended to improve the data set, but which are not considered critical to the understanding and use of the data. Major liens are those severe enough to render the data set (or increment, in the case of dynamic data sets) to be unsuitable for scientific research.

logical identifier (LID): An identifier which identifies the set of all versions of an object.

metadata: Data about data — for example, a 'description object' contains information (metadata) about an 'object.'

product: One or more tagged objects (digital, non-digital, or both) grouped together and having a single PDS-unique identifier. In the PDS4 implementation, the descriptions are combined into a single XML label. Products are the smallest addressable granular unit in the PDS holdings.

10 References

[1] JPL: Basics of Space Flight, Chapter 7, https://solarsystem.nasa.gov/basics/bsf7-1.php

[2] NASA , 2014 Discovery Proposal Evaluation Plan. http://discovery.larc.nasa.gov/discovery/pdf_files/Discovery_2014_AO_Eval_Plan_20150817_public_version.pdf