Planetary Data System

Proposer's Archiving Guide (PAG)

June 15, 2003

Version 1.0

Jet Propulsion Laboratory
Pasadena, California

JPL D-26359

Planetary Data System

Proposer's Archiving Guide (PAG)

June 15, 2003

Version 1.0

Approved:

____________________________________ _____________
Laverne Hall Date
PDS Project Manager

____________________________________ _____________
Reta Beebe Date
PDS Project Scientist

____________________________________ _____________
Bill Knopf Date
PDS Program Executive

JPL

Jet Propulsion Laboratory
Pasadena, California

JPL D-26359

CHANGE LOG

Revision	Date	Description	Author
Start Draft	April 2, 2003	Initial outline	R. Joyner
Draft Release 1	April 17, 2003	First Draft	V. Henderson
Draft Release 2	April 22, 2003	Add Sections 1.1, 1.2; reworked 2.0, edited document throughout to reduce level of detail	V. Henderson
Draft Release 3	May 8, 2003	Added Section 4 Cost Model	V. Henderson
Draft Release 4	May 9, 2003	Change "project" to "mission". Added text to Sections 2.2, 3.1, 3.2; added acronym list;	V. Henderson
Draft Release 5a	16 May 2003	Section 1	R.A. Simpson
Draft Release 5b	18 May 2003	All	R.A. Simpson
Draft Release 6	25 May 2002	All sections – using submittals from PAG working group	R. Joyner
Draft Release 1.7	2003-06-01	All sections - more inputs from PAG WG	Dick Simpson
Draft Release 1.8	2003-06-05	Ruthless reduction	M. Gordon
Draft Release 1.9	2003-06-11	All sections - more inputs from PAG WG	M. Gordon
Draft Release 1.10	2003-06-14	Penultimate clean up	M. Gordon
Initial Release 1.0	2003-06-15	All sections; added URLs	R. Joyner

TABLE OF CONTENTS

1.0 INTRODUCTION

1.1 Purpose

The purpose of this document is to provide basic information on the archiving process (including PDS expectations) to scientists proposing for NASA planetary programs. THE PDS hopes that better understanding of archiving at the earliest stage will help proposers address key issues, integrate archiving with the data production pipeline, and allocate adequate resources for archiving. The result should be higher quality archives produced more efficiently at lower overall cost.

1.2 Scope

This document addresses the following topics broadly, but with enough detail that scientists should be able to draft the archiving section of a planetary proposal. A generic 'mission' context is assumed since the majority of data (both in volume and complexity) flows from NASA missions.

- Overview of the archiving process, including goals
- Elements of a good archive
- Organization and structure of the Planetary Data System
- Mission archiving roles and activities
- Steps in the archive process
- Costing the archive activity
- Typical deliverables

Appendix C provides links to additional resources and sample documents .

1.3 Audience

This document is written primarily for scientists responding to NRA's and AO's for NASA planetary missions. Others (proposers for individual instruments, proposers to data analysis programs, foreign investigators seeking to archive with PDS) should find the material useful, after appropriate adjustments. The expected audience for this document includes:

- Proposers to NASA mission-scale opportunities
- Proposers for investigations to supply instruments on NASA missions
- Proposers to NASA data analysis programs
- Scientists and managers in non-U.S. missions
- Scientists interested in restoring and archiving old data sets
- Users of data in the existing PDS holdings
- NASA managers at various levels
- Archivists and data management professionals

1.4 Applicable Documents

[1] Planetary Data System Standards Reference, October 15, 2002, Version 3.5, JPL D-7669, Part 2. (http://pds.jpl.nasa.gov/documents/sr/)

[2] Planetary Science Data Dictionary Document, August 28, 2002, Planetary Data System, JPL D-7116, Rev. E. (http://pds.jpl.nasa.gov/documents/psdd/psdd.pdf)

2.0 OVERVIEW OF THE ARCHIVING PROCESS

The PDS archives and distributes documented data to the planetary science community. It was established in response to requests for increased availability of planetary data and to concerns that data from past NASA missions were degrading to the point of becoming unusable for analysis. PDS functions include both ingestion and distribution of data at 'nodes' specializing in certain scientific disciplines and/or technical skills. A central catalog provides high level information on node holdings.

In order to fulfill its charter, the PDS requires that data it handles be understandable, be in formats that future scientists will find easy to use, and follow standards for organization and content that facilitate machine-assisted correlative science across missions and science disciplines. The PDS requires that submitted data meet published standards regarding format, content, and documentation [1, 2]. A range of formats has been defined, but 'tables' and 'images' account for the majority of products. Data products are organized logically into 'data sets,' which are sometimes organized into 'data set collections'. Data may be written to 'volumes' and 'volume sets' on physical media such as CD-ROMs and DVDs.

At an early stage 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. All archive submissions to the PDS are peer-reviewed by scientists and data engineers to ensure that PDS standards have been met and that the archive is complete and the data are useful. This usually occurs at several stages for mission archiving.

Once data are fully integrated into the PDS, they may be retrieved through electronic queries over the internet, or by special request on physical media such as CD-ROM. The PDS provides copies of all accepted data sets to the National Space Science Data Center (NSSDC), which serves as both the PDS 'deep archive' and the official distribution point for requests originating outside the planetary research community.

3.0 ELEMENTS OF A GOOD ARCHIVE

3.1 Overview

A useful mission archive, as required by most NRA's and AO's, includes raw data formatted from each instrument, data calibrated in physical units, and derived products based on further processing of the data and/or combinations of different data such as maps, overlays, and comparative tables. The archive contains sufficient documentation of the mission, the instruments, and the calibrations that a scientist of a future generation can intelligently use the data and, if appropriate, even recalibrate the data. The archive includes complete information about the geometry relevant to the mission (e.g., spacecraft position and orientation relative to the target). It includes indices to help in searching the data, catalog files to enable ingestion of meta-data into the PDS database, supporting documentation, and any other information relevant to understanding or interpreting the data. Finally, the archive uses a straightforward organization and a small number of widely recognized, non-proprietary formats.

An instrument archive designed in advance and fed by a data processing pipeline contains 'standard' products, usually from the single instrument, but possibly from an integrated payload of several instruments. As the mission evolves, new 'special' products (combinations of data from multiple sources, the results of analyses not originally planned, etc.) may be defined. The PDS encourages inclusion of these additional special products in the archive; however, the decision to do so rests with the mission (or instrument) and often depends on availability of resources.

3.2 Instrument Data Products

Appendix B lists the NASA and the CODMAC (Committee On Data Management and Computation) data levels. Missions are normally expected to archive both raw data (formatted appropriately for each instrument) and data fully calibrated in physical units. These usually correspond to CODMAC levels 2 and 3, although there may be cases in which an unusual instrument calibration is inherently irreversible, thus leading to CODMAC level 4. Where no investigation team has been selected, e.g. for radio science, the mission should still consider archiving the data from all major subsystem operations if there is a potential for any scientific use. Most missions transfer CODMAC level 1 data to an entity in the implementing organization and/or some part of the PDS to be "saved", but these are not formally part of the archive.

The archive should include a complete description of the operation of each instrument and of the calibration procedure and all files needed to carry out the calibration. The PDS also requires a short, descriptive file about each instrument for cataloging in its central database. Formats of the data, both raw and calibrated, should be simple, non-proprietary, and widely used—formats that would be familiar to most in the relevant segment of the research community. Each data file has a label, normally (but not always) a detached label in a separate file, that describes the data file itself in a structured keyword=value format. This label is intended to be read by humans and by data analysis software.

3.3 Other Mission 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 coordinate systems on the target, a history of all significant spacecraft events, and other housekeeping data (such as temperatures and power levels) that might be useful in understanding the behavior of instruments.

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.

Normally, radiometric tracking data should be archived even if there is no "instrument team" for radio science.

The mission must also provide suitable catalog files describing the overall mission, the spacecraft, and the instrument(s) so that future scientists will be able to use the data intelligently.

Index tables must be provided to assist in locating individual data products in the archive.

3.4 Higher level products

AOs and NRAs may also require, and the PDS encourages, the mission to archive higher-level products, such as those at CODMAC levels 4 and above. These might include mosaics, maps, data resampled to a common interval, fits to model results, or inversions. Higher level products may come from a single instrument or from combining the results from several instruments.

3.5 Documentation

Second in importance only to the data products themselves is the documentation which accompanies them in the archive. The documentation takes several forms. Every individual file must be described in its associated label. For some files, the label is included at the beginning of the file. However, for many files, especially data products, it is more appropriate to include the label as a separate file. Labels are intended to be readable by humans as well as by data analysis software.

3.5.1 Data product Documentation

Each product type destined for archive should be defined in an Interface Control Document (ICD) — sometimes called a Software Interface Specification (SIS) — which describes its format (at the bit level if necessary) and content, including processing history and definition of all fields. The ICD is a "living" document in the sense that it is continually revised to reflect changes in the data product. The most current version of the ICD should be included in the archive.

3.5.2 Archive Volume Documentation

It is common to document the volume structure in an Archive Volume ICD. The Archive Volume ICD may be combined with the individual data product ICD's as separate chapters in a single document. In either case, the resulting document(s) should be included in the archive.

3.5.3 Calibration Documentation

Each instrument team should include a description of the calibration procedure (or algorithms) and all necessary data files so that future users can reproduce the calibration. The calibration procedure should be appropriate to the processing level of the data in the archive. Sometimes it is useful to include computer source code as documentation. However, providing source code alone, expecting all future scientists to use it, is not considered an acceptable response to the calibration requirement because software is almost always platform dependent and no software can be guaranteed stable over extended periods of time regardless of the platform.

3.5.4 Catalog and General Documentation Files

The mission team is usually responsible for drafting and maintaining descriptions of the spacecraft and the mission itself, along with a list of references. Each instrument team is usually responsible for drafting and maintaining descriptions of its instrument (the instrument, its modes, and the scientific objectives of the investigation) and its data sets (including coverage and quality). These descriptions are included in each instrument archive. They are also entered into the PDS online catalog to be available for searching, and are therefore known as the catalog files.

Additional documentation such as journal submissions describing the instrument and investigation, technical reports, analysis algorithms, operations plans, and other material may also be included in the archive. Software may be submitted as either documentation or operational modules with the caveats noted above under ‘Calibration Documentation’.

3.6 Archive Organization

As noted above, data products are organized logically into data sets but are physically written to volumes; large data sets may require multiple volumes. Mission data are usually delivered incrementally to the PDS every 3-6 months—sometimes negotiated to be 'electronic' rather than physical volumes. A typical volume structure is shown in Figure 3-1.

```
root of the archival volume
   |
   |- AAREADME.TXT      text file describing volume
   |- ERRATA.TXT        cumulative errors and notes for volume set
   |- VOLDESC.CAT       structured text describing volume
   |
   |- CALIB             directory for calibration files
   |   |- CALINFO.TXT   text file describing directory contents
   |   |- CALDATA1.TAB  first calibration file
   |   |- CALDATA1.LBL  label for CALDATA1.TAB
   |   |-     ...
   |
   |- CATALOG           directory of 'catalog' files
   |   |- CATINFO.TXT   text file describing directory contents
   |   |- MISSION.CAT   text file describing mission
   |   |- INSTHOST.CAT  text file describing instrument host
   |   |- INST.CAT      text file describing instrument
   |   |- DATASET.CAT   text file describing data set
   |   |- PERSON.CAT    text file listing contributors to archive
   |   `- REF.CAT       text file of references cited in other files
   |
   |- DOCUMENT          directory containing documents
   |   |- DOCINFO.TXT   text file describing contents of directory
   |   |- ICD.TXT       volume and data product ICD
   |   |- DOCUMNT2.TXT  second document file
   |   |-     ...
   |
   |- GEOMETRY          directory containing geometry files
   |   |- GEOMINFO.TXT  text file describing contents of directory
   |   |- GEOM1.DAT     first geometry file
   |   |- GEOM1.LBL     label for GEOM1.DAT
   |   |-     ...
   |
   |- INDEX             directory with tables of contents files
   |   |- INDXINFO.TXT  text file describing directory contents
   |   |- INDEX.TAB     table of contents for the volume
   |   |- INDEX.LBL     label for INDEX.TAB
   |   |- CUMINDEX.TAB  cumulative table of contents for volume set
   |   `- CUMINDEX.LBL  label for CUMINDEX.TAB
   |
   `- DATA              directory for data
       |- DATAFIL1.TAB  first data table
       |- DATAFIL1.LBL  label for DATAFIL1.TAB
       |- DATAFIL2.IMG  first image file
       |- DATAFIL2.LBL  label for DATAFIL2.IMG
       |-     ...
```

Figure 3-1: Typical volume structure for a data set spanning multiple volumes. The CATALOG directory contains structured documentation on the data set, the associated instrument(s), the spacecraft or other host, the mission, and other matters; these files become part of the PDS central node catalog. The GEOMETRY directory contains ephemerides, orientation data, and other geometry data. The INDEX directory contains tables of contents for both the volume and volume set. The DATA directory is often divided into subdirectories grouping data by type, source, or time

4.0 THE PLANETARY DATA SYSTEM (PDS)

This section briefly describes the organization of the PDS, some of its responsibilities, and how those are distributed within the system.

4.1 Roles and Responsibilities (within the PDS)

The role of the PDS is to ensure that future generations of scientists can make good use of data returned from today's NASA missions. The PDS comprises several operational units or 'nodes.' Each node is staffed and operated independently at a university or government research center, and curates data within a discipline of planetary science and/or provides specialized technical expertise to the PDS.

4.1.1 Central Node

The PDS Central Node manages and coordinates the PDS. For purposes of this document its most important roles are as the initial PDS contact to new missions, in ensuring that interfaces developed later at the discipline node level function smoothly, and as a source of tools and standards for constructing and validating archives.

4.1.2 Discipline Nodes

Each discipline node is responsible for data in a specific subject area or specialty; each maintains the PDS data holdings in that specialty. At least one discipline node will work closely with the mission team throughout archive development.

This is a brief summary of the current nodes.

PDS Central Node, located at the NASA Jet Propulsion Laboratory (Pasadena, CA), manages and coordinates PDS. (http://pds.jpl.nasa.gov/)

Planetary Atmospheres Node, located at New Mexico State University (Las Cruces, NM), concentrating on planetary atmospheres and including the gas giant planets. (http://pds-atmospheres.nmsu.edu/)

Geosciences Node, located at Washington University (St. Louis, MO), specializing in the surfaces and interiors of terrestrial planets. (http://pds-geosciences.wustl.edu/)

Planetary Plasma Interactions Node, located at the University of California, Los Angeles, concerned with fields and particles, plasmas, and interactions with the interplanetary medium. (http://pds-ppi.igpp.ucla.edu/)

Planetary Rings Node, located at the NASA Ames Research Center (Moffett Field, CA), and concentrating on planetary ring systems and dynamical interactions with inner moons. (http://pds-rings.arc.nasa.gov/)

Small Bodies Node, located at the University of Maryland (College Park, MD), and concerned with comets, asteroids, Pluto, trans-Neptunian objects, and interplanetary dust. (http://pds-smallbodies.astro.umd.edu/)

Imaging Node, located at the U.S. Geological Survey (Flagstaff, AZ) and Jet Propulsion Laboratory (Pasadena, CA), specializing in digital imaging and image processing. Imaging has also accepted responsibility for large icy satellites. (http://pds-imaging.jpl.nasa.gov/)

Navigation and Ancillary Information Facility (NAIF), located at the Jet Propulsion Laboratory (Pasadena, CA), specializing in observational geometry (ephemerides and orientations, including the SPICE information system) and ancillary information (e.g., spacecraft engineering data). (http://pds-naif.jpl.nasa.gov/)

4.1.3 PDS Roles in Mission Archiving

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

* Mission contacts
* Instrument / Science Team contacts
* Discipline Node contacts
* Central Node contacts

A new mission is usually assigned to a single 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 (including potential data contributors without specific instruments, such as Interdisciplinary Scientists) through other discipline nodes if appropriate. The lead node works with the mission to ensure that archive plans are developed, designs meet standards and facilitate cross-instrument and cross-mission comparisons, and archive production and validation run smoothly.

5.0 MISSION ARCHIVING ROLES AND ACTIVITIES

The PDS must establish contacts with a mission at several levels. Because missions themselves vary in organizational complexity, no single model is adequate to describe these relationships. However, certain key functions and responsibilities are needed in all cases. This section describes the archiving organization and contact points which recur in many missions. The discussion is meant to be illustrative, not exclusionary.

5.1 Organizing the Archiving Activity

It is at this point that missions differ substantially, depending on the number of different instruments and the organization of the entire mission team. Small missions with few instruments might have only the Principal Investigator, the Archive Scientist, and some technical support to deal with archiving. The Archive Scientist might be in charge of instrument calibration, data processing, and archive production.

Complex missions will have separate instrument teams for each instrument and usually delegate archiving responsibilities to the instrument teams. These teams may deliver their archival products directly to the PDS or may route them through a mission data center. Missions for which there are separate instrument teams will almost always find it valuable to create a Data Archive Working Group chaired by the Mission Archive Scientist.

5.2 Mission Archiving Roles

5.2.1 Principal Investigator (PI)

Recent NRAs and AOs generally make it the responsibility of the Principal Investigator and the investigation team to deliver archival data products to PDS that meet PDS standards. While the PI remains responsible, the authority for carrying out the work may be delegated, either in small part or almost entirely, to one or more specific members of the investigation team.

Coordination is required between the top levels of the mission and the PDS. The mission must issue a Data Management and Archiving Plan (DMAP). This may be written as two separate documents, a Data Management Plan (DMP) and an Archiving Plan (AP). In either case, these documents define the scope of the archive and thus the costs of archiving, both for the mission in preparing the archival products and for PDS in reviewing the data, archiving the data, and subsequently distributing it to the community. Therefore, the PI, the PDS Project Manager, and the node manager of the lead node must agree on the content of the archive as described in these documents.

The PI is responsible for ensuring an adequate budget for archiving, including resolution of liens from any peer review process planned after the final archive delivery.

5.2.2 Mission Archive Scientist

It is generally more efficient for the mission/investigative team if a single person is identified as the Mission Archive Scientist. This would not normally be a full-time position except perhaps on very large, complex missions. This individual oversees the archiving activities of the entire mission and ideally is very familiar with PDS practices. The Mission Archive Scientist is normally the primary contact between the mission and the PDS lead node and serves as the mission's clearing house for archiving information.

Potential tasks and responsibilities for the Mission Archive Scientist generally include:

- Prepare the archive portions of the Data Management and Archiving Plan.

- Coordinate the preparation of the Interface Control Document that describes in detail the content of the archive.
- Prepare the descriptive text about the mission, instrument host (*e.g.*, spacecraft), and reference catalog files for the PDS database.
- Ensure the timely development by designated mission team members of additional archive documents (e.g. individual instrument ICDs, instrument catalog files, etc.).
- Ensure that any copyright issues have been addressed if previously published material is to be included in the archive.
- Ensure compatibility of formats and definitions from one instrument to another.
- Coordinate delivery schedules for the archival products from different instruments.
- 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 for personnel from PDS or from other organizations to present tutorials.
- Arrange any desired technical workshops for the mission.

5.2.3 Archive Preparers

Generally, the mission team will identify one or perhaps two archive preparer(s) per instrument. 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, but frequently they are split between two people. Early and continued interaction with PDS personnel, both PDS discipline scientists and PDS data engineers, will be invaluable in optimizing the efficiency of creating the archive.

Potential tasks and responsibilities for archive preparers include:

- Define instrument archival products.
- Design the details of each data product (file naming, keywords, file formats, etc.).
- Ensure the scientific validity of the archival products.
- Prepare the archive portions of the instrument Interface Control Document that describes all these details.
- Develop the INST.CAT and DATASET.CAT files, and assist in writing the PERSON.CAT file.

- Ensure that the products and volumes meet PDS standards.
- Production and transmission to PDS of the archival volumes(s).

6.0 STEPS IN THE ARCHIVE PROCESS

This section lists the usual steps in a creating a successful mission archive, including the key milestones and deliverables. Terminology for the phases and the various reviews may vary from one program to another (*e.g.* between Discovery and New Frontiers) or even from one implementing organization to another. Figure 6-1, at the end of this section, depicts the same information in chart form.

6.1 Pre-proposal Briefing

Prior to proposal deadlines for NRAs and AO's, 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 *Proposer's Archiving Guide* (this document) will be provided at the briefing.

6.2 Proposal Phase

Typical AO's require that proposals include a discussion of products to be delivered to PDS and that budgets include appropriate funding for this activity. During the proposal phase, PDS will make available sample archiving documents and a tool for estimating archival costs. PDS personnel are also available to answer archive design questions from proposal teams on a confidential basis.

6.3 Concept Study Report (CSR)

In the Discovery and New Frontiers programs, several original proposals may be selected to go into an advanced round of conceptualization (roughly equivalent to Phase A in NASA’s terminology for center-led missions) and leading to a Concept Study Report (CSR), review, and down-selection. In some cases the science section of the proposal, which contains the discussion of archiving, is meant to remain unchanged in this phase. However, in some instances changes may be allowed in order to strengthen the proposal. During this phase, PDS will provide a detailed PDS orientation to any CSR team requesting it. Proposal teams are welcome to contact PDS, either the central node or an appropriate discipline node, with questions on a confidential basis. Since much of the content of the archive plan must be thought through in order to develop a realistic CSR budget, sometimes a preliminary archive plan is included as an appendix to the CSR.

6.4 Phase B

For relatively small, PI lead missions, a Memorandum of Understanding (MOU) may be appropriate. If so, a Memorandum of Understanding (MOU) defining/delineating roles and mutual responsibilities regarding archiving and distribution of mission data should be negotiated between the mission and PDS early in Phase B. (Note that while an MOU may be optional for small missions, it is essential for large center lead missions and should be negotiated during Phase A). PDS will assign a discipline node as the primary contact with the mission (referred to as the 'lead node' for the mission). In many cases, the choice of lead node will be obvious; proposers may have even specified a lead node in their proposal and/or CSR. The lead node and a data engineer from the central node will work with the mission to complete the archive portion of the Data Management and Archiving Plan (DMAP). The DMAP will contain sufficient detail about the data flow, archival products, and procedures for generating the archive that everyone understands the responsibilities. Additionally the DMAP 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. The lead node and the mission will decide whether other discipline nodes should be enlisted for handling particular data sets. The lead node will advise the mission, but not set requirements, on how to organize the archiving for the mission. This is likely to vary depending on the complexity of the mission. A complete draft of the Data Management and Archiving Plan is required at the Preliminary Design Review (PDR).

6.5 Phase C/D – Detailed Design, Construction, and Test

During this phase archive generation is designed as an integral part of the data processing system. The DMAP must be signed by all parties by the time of the Critical Design Review (CDR). Some missions may review the operations (Ops) and ground data system (GDS) separately from the flight system in which case the signed DMAP can be provided for the Ops/GDS CDR.

By the time of instrument tests, the archive must be designed in sufficient detail that any test and calibration data likely to be useful in flight can be archived without significant additional work. Details of attitude telemetry should also be developed to a similar level. An Interface Control Document (ICD) should be outlined at an early stage and filled in as details are decided in the hardware design. The ICD describes in detail the content of the archive, including definitions of all parameters, calibration procedures, and details of formats—to the bit level, if appropriate. A complete, but probably not final, ICD must be available at the Mission Readiness Review (MRR); details in the ICD may change as analysis procedures and data formats are refined once operations begin.

During Phase C/D, close contact between the mission and lead node will minimize possible production and delivery problems. A review of the processing pipeline and sample archival output should be conducted near the end of Phase C/D or early in Phase E. PDS can provide assistance for peer reviews carried out by the mission, and it can organize external peer reviews. For a long-duration mission with a steady flow of data, pipeline and sample data reviews are an important mechanism for minimizing later problems that might require reprocessing of large volumes of data.

6.6 Phase E (Operations)

Details of interactions between PDS and missions will vary considerably during the operations phase, depending on when data deliveries are scheduled and whether there is a fixed pipeline delivering large volumes of data. There are four key "events" in Phase E, some or all of which may be repeated periodically if the operations phase is a long one. These are 1) deliveries to PDS, 2) updates to the ICD reflecting the archive as actually produced, 3) peer review of the archive, and 4) resolution of all liens identified in the peer review.

Timing of data deliveries will be determined by guidelines specified in the AO, by specific agreement between the mission and NASA as part of the proposal and selection process, and/or by negotiation between the mission and PDS. There may be a single delivery of all data at the end of a short mission or there may be multiple, incremental deliveries for a mission with a long operations phase. There may be a continuously operating pipeline that does not change across multiple deliveries, or the details of the products may vary from one delivery to another. The deliveries may all come from a single project-wide data center, or the deliveries may come separately from each of several instrument teams. These choices will have been made and documented in the mission DMAP.

A peer review should be conducted of the earliest sample data from each processing pipeline, thus ensuring a minimum of late-arising problems. Depending on the detailed organization of the mission, these reviews might occur late in Phase C/D or early in Phase E. Each data delivery must be accompanied by the current, complete ICD for that delivery. PDS will organize peer reviews and these will involve both technical specialists on archiving and scientific peers who are not connected with either PDS or the mission (much like referees of journal articles). The mission is responsible for addressing all liens identified in the peer reviews, just as an author is responsible for addressing all comments of a journal referee. As with journal refereeing, this does not always mean doing everything suggested by the review panel, but it does mean doing many of the things and explaining why other requests for change were not accepted. The scientific peers are often the most severe critics of an archive. The final data submission should occur prior to the end of the mission. The mission must hold sufficient reserves at the end of the mission to address any liens from the final peer review of the delivered archive. Minimizing problems at this very late stage is the strongest reason for working closely with PDS at an early stage and using early peer reviews of sample data.

6.7 Major Documents and Reviews by Phase

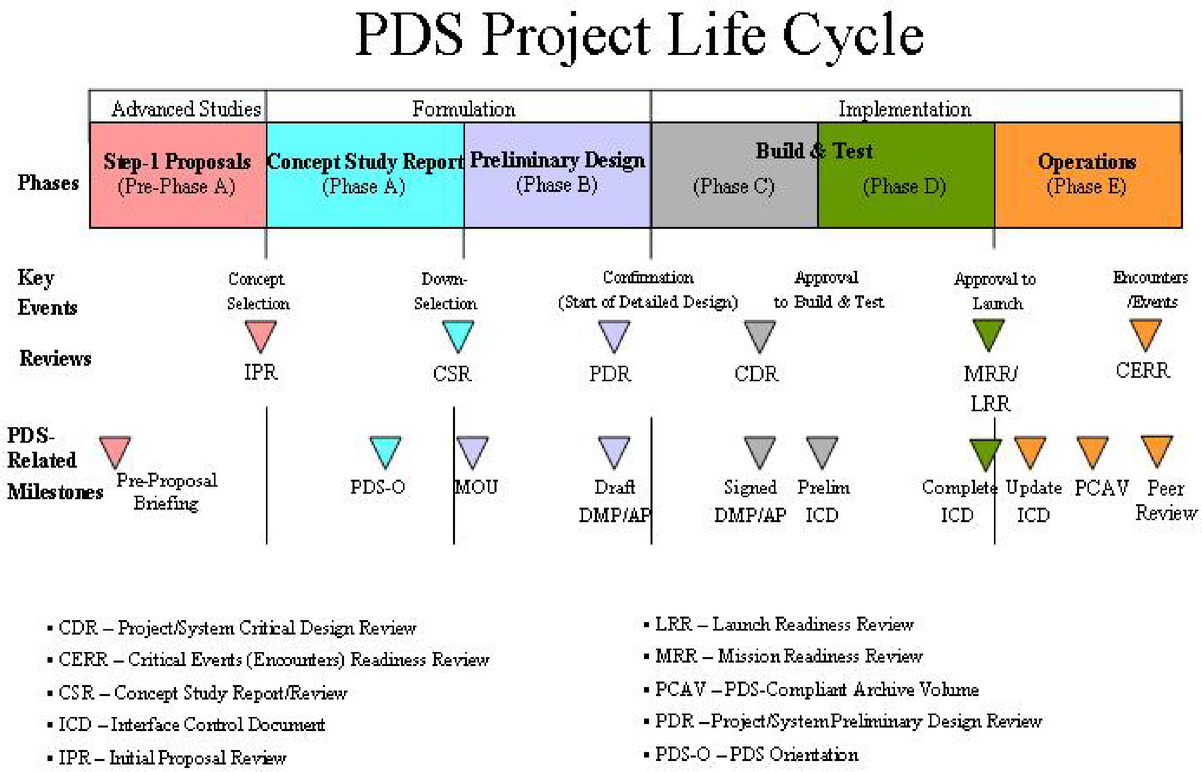

Figure 6-1. PDS Project Life Cycle

7.0 DELIVERABLES EXPECTED BY THE PDS

The following deliverables are expected by the PDS to establish coordination and communication channels, determine roles and responsibilities, and to define the scope of an archive; all may become components of the archive itself.

7.1 Memo of Understanding (MOU)

The MOU is a short document negotiated between a new mission and the PDS defining/delineating roles and mutual responsibilities regarding archiving and distribution of mission data. The MOU also contains a list of documents to be produced and the parties responsible for producing them. The MOU is produced early in Phase B. See PAG web page for examples. (http://pds.jpl.nasa.gov/documents/pag/)

7.2 Data Management and Archiving Plan (DMAP)

The DMAP is a general description of mission data processing, cataloging, and communications and a more detailed description of the generation, validation, and transfer of the mission data archive to the PDS. The DMAP may be delivered as a single document or as separate Data Management Plan (DMP) and Archiving Plan (AP). Drafts of these are prepared during Phase B and finalized in Phase C. See PAG web page for examples. (http://pds.jpl.nasa.gov/documents/pag/)

7.3 Interface Control Document (ICD)

The mission will prepare one or more ICD(s) which describe the content and organization of each archive volume and the format, content, processing history, and labeling of each data product. Draft ICDs are prepared in Phase C, reviewed in Phase D and 'finalized' in Phase E. Note that ICDs are living documents; consequently revisions may be necessary to the 'finalized' versions. See PAG web page for examples. (http://pds.jpl.nasa.gov/documents/pag/)

7.4 Catalog Files

Files in the CATALOG directory are delivered on schedules given in the Archive Plan. Typically mission and spacecraft descriptions (in the files MISSION.CAT and INSTHOST.CAT) are drafted and maintained by mission personnel, while instrument, data set, and personnel information (in the files INST.CAT, DATASET.CAT, and PERSON.CAT) are drafted and maintained by the respective instrument teams. Instrument teams contribute to a master reference list (REF.CAT), which is maintained by the mission. The information in these files becomes part of the PDS Catalog.

Draft versions of catalog files are delivered during Phase C for purposes of distribution. Initial versions of catalog files must be delivered no later than early Phase E. The catalog files will be updated throughout the mission with the final updated versions submitted to the PDS at the end of the mission. See PAG web page for examples. (http://pds.jpl.nasa.gov/documents/pag/)

7.5 Lien Resolutions

Resolutions to all problems identified during peer reviews are required before the PDS accepts archive volumes.

7.6 PDS Compliant Archive Volumes

The mission will deliver three copies of each archive volume to the PDS, on PDS compliant archival media, containing all data products designated in the Archive Plan, compliant with PDS standards, and on schedules defined in the archive plan.

Data products are delivered during Phase E.

8.0 COSTING THE ARCHIVE ACTIVITY

Satisfactorily fulfilling archive requirements embodied in NRAs and AOs can absorb a substantial amount of time and funds from a mission's budget. However, how the mission organizes its archiving efforts (designation of a Mission Archive Scientist, formation of a Data Archive Working Group, etc.) and how the mission designs the archive and its generation can significantly reduce the impact of archiving on the mission team. Experience shows that including archive product generation as an integral part of the data processing pipeline is likely to produce the greatest single reduction in archiving costs for the mission in terms of both time and money. Another opportunity for significant savings arises from early archive design review and testing including the production and review of sample data products. Early and frequent interaction between the mission archiving activities and the PDS can further reduce the overall archiving costs.

The PDS has developed a first-order model to estimate costs incurred by a mission or instrument team in designing and delivering a PDS-compliant archive. Although the model is based on recent archiving experience, it is not a precise predictor of costs and should be used as a guide, not as a rigorous prediction tool. It has been incorporated into a Microsoft Excel spreadsheet, which will run on either a Windows or Macintosh platform. The Mission Cost algorithm and tool are available for download at:

http://pds.jpl.nasa.gov/documents/pag/

APPENDIX A ACRONYMS

The table below lists acronyms and abbreviations used in this document.

AO	Announcement of Opportunity
AP	Archive Plan
CDR	Critical Design Review
CD-ROM	Compact Disc—Read-only Memory
CODMAC	Committee On Data Management And Computation
CSR	Concept Study Report
DMAP	Data Management and Archiving Plan
DMP	Data Management Plan
DVD	Digital Versatile Disc
FTE	Full Time Employee or Full Time Equivalent
GDS	Ground Data System
ICD	Interface Control Document
MRR	Mission Readiness Review
NAIF	Navigation and Ancillary Information
NASA	National Aeronautics and Space Administration
MOU	Memorandum of Understanding
NRA	NASA Research Announcement
NSSDC	National Space Science Data Center
Ops	Operations
PDR	Preliminary Design Review
PDS	Planetary Data System
PI	Principal Investigator
SIS	Software Interface Specification
SPICE	Spacecraft ephemeris, Planet/satellite ephemeris, Instrument information, Camera orientation, Event information.
TBD	To be determined
URL	Uniform Resource Locator
USGS	United States Geological Survey

Table A-1. Acronyms and abbreviations

APPENDIX B NASA AND CODMAC PROCESSING LEVELS

NASA	CODMAC	Description
Packet data	Raw - Level 1	Telemetry data stream as received at the ground station, with science and engineering data embedded.
Level-0	Edited - Level 2	Instrument science data (e.g., raw voltages, counts) at full resolution, time ordered, with duplicates and transmission errors removed.
Level 1-A	Calibrated - Level 3	Level 0 data that have been located in space and may have been transformed (e.g., calibrated, rearranged) in a reversible manner and packaged with needed ancillary and auxiliary data (e.g., radiances with the calibration equations applied).
Level 1-B	Resampled - Level 4	Irreversibly transformed (e.g., resampled, remapped, calibrated) values of the instrument measurements (e.g., radiances, magnetic field strength).
Level 1-C	Derived - Level 5	Level 1A or 1B data that have been resampled and mapped onto uniform space-time grids. The data are calibrated (i.e., radiometrically corrected) and may have additional corrections applied (e.g., terrain correction).
Level 2	Derived - Level 5	Geophysical parameters, generally derived from Level 1 data, and located in space and time commensurate with instrument location, pointing, and sampling.
Level 3	Derived - Level 5	Geophysical parameters mapped onto uniform space-time grids.

Table B-1. NASA and CODMAC Definitions of Processing Levels for Science Data Sets

APPENDIX C ONLINE RESOURCES

The following resources may be of interest during proposal preparation and are available electronically at either the designated URL or can be located at the Proposers Archiving Guide (PAG) web page:

http://pds.jpl.nasa.gov/documents/pag/

1. PDS Home Page (http://pds.jpl.nasa.gov/)

2. Web-based (quick-start introduction) to archiving data with PDS (http://pds.jpl.nasa.gov/documents/qs/)

3. PDS Standards Reference (JPL Document D-7669, Part 2) -Specific PDS data preparation standards for archive quality data sets. (http://pds.jpl.nasa.gov/documents/sr/)

4. PDS Data Dictionary ...- List of PDS-recognized keywords (and acceptable values where restrictions apply) (http://pds.jpl.nasa.gov/documents/psdd/psdd.pdf)

5. Proposer's Archiving Guide Version 1.0 (June 15, 2003) (This document)

6. Mission Archiving Cost Analysis Model / Algorithm

   - Cost Analysis Algorithm (MSWord)
   - Cost Analysis Tool (XLS)

7. Example Archive Documents:

   - MOU – Example Memorandum of Understanding (MRO)
   - DMAP – Example Data Management and Archiving Plan (Messenger & Deep Impact)
   - DMP – Example Data Management Plan, the first part of the DMAP when two separate documents are prepared (TBD)
   - AP – Example Archiving Plan, the second part of the DMAP when two separate documents are prepared (draft Cassini)
   - ICD/SIS – Example document describing volume and data product structure, content, and processing history (draft Huygens-SSP; Data Product SIS template; Archive Volume SIS template)

8. Example Catalog Files:

   - MISSION.CAT (Voyager)

- INSTHOST.CAT (Voyager 1)
- INST.CAT (Voyager / RSS-VG1S)
- DATASET.CAT (Voyager 2 Triton Radio Occultation Reduced Data V1.0)
- PERSON.CAT
- REF.CAT
- TARGET.CAT

9. Example Products and Labels

    - Example IMAGE data product and data product label (Venus)
    - Example TABLE data product and data product label (INDEX files)

10. Detailed Archiving Guide (TBD)