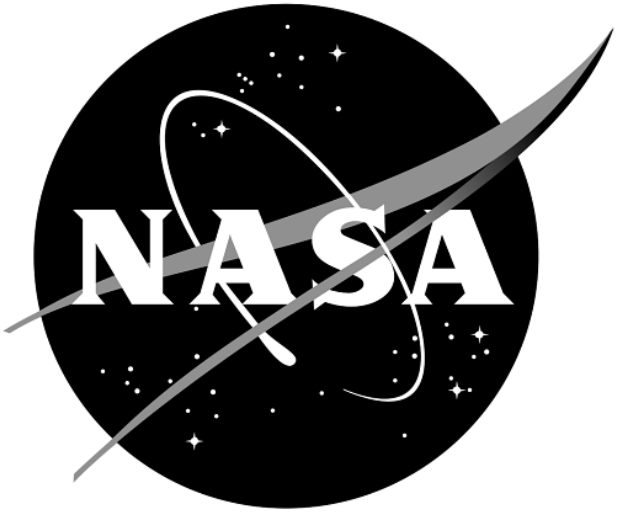
[redacted]

Plan Document

NASA Planetary Data System

PDS 2010 Project Transition Plan



Change Log

Revision	Date	Description	Author
Draft		Initial draft release.	

Contents

1 Introduction

For over fifteen years, the Planetary Data System (PDS) has been NASA's official data system for archiving and distribution of data from planetary exploration missions. It has been a leader in defining data standards, working with missions and instrument teams, and developing data system technologies. The PDS has been instrumental in changing the scientific culture by working with the planetary science community to publicly release and peer review the data it captures. It has also been used as a model by other science data systems interested in establishing distributed scientific networks organized by independent discipline nodes at facilities that are doing leading-edge scientific research.

While PDS has been a leader in developing and exploiting new technologies and ideas, an increasing workload and substantial increases in the volume of delivered data are now threatening the system's ability to accomplish its primary missions of both archiving planetary science data and distributing it to working scientists. PDS identified these challenges in its Roadmap published in 2006. In addition to these challenges, the ten year Roadmap outlined several goals including improving the PDS data standards, increasing user services by leveraging newer technologies and technical standards, and re-architecting PDS to ensure efficient operations of the system while supporting the increasing demands on PDS by both the data providers and end users.

In response to these challenges and goals, PDS has developed a plan for the next generation called "PDS 2010". The vision for PDS 2010, as defined by the PDS Management Council at its April 2008 meeting, includes:

- Simplified, but rigorous, archiving standards that are consistent, easy to learn, and easy to use
- Adaptable tools for designing archives, preparing data, and delivering the results efficiently to PDS
- On-line services allowing users to access and transform data quickly from anywhere in the system
- A highly reliable, scalable computing infrastructure that protects the integrity of data, links the nodes into an integrated data system, and provides the best service to both data providers and users

1.1 Purpose

The purpose of this transition planning is to layout the tasks and activities that need to take place to efficiently move the PDS 2010 system from the pilot environment to the production, operations

and maintenance environment. The objective is to ensure minimal impact on continuity of system operations.

1.2 Scope

The scope of the PDS 2010 Project involves upgrades to the following areas:

1. Data Standards
2. Software tools to support archive preparation
3. Software services to support ingestion and distribution
4. Archive processes to improve efficiency end-to-end of PDS
5. Software applications for accessing PDS data

This transition plan includes strategy and concept of system transition and data migration. It identifies stakeholders that are affected by the transition. It captures what are the tasks and activities required of the stakeholders to support the transition, and when they should occur.

1.3 Document Revision

Revisions of this document will be held in the PDS Engineering Node website through the use of its document history functionality. Previous versions of this document can be accessed through the use of that tool.

1.4 Applicable Documents

1.4.1 Controlling Documents

[1] Planetary Data System Strategic Roadmap 2006 - 2016, February 2006.

[2] Planetary Data System Level 1, 2 and 3 Requirements, August 2006.

1.4.2 Referenced Documents

[3] PDS 2010 Project Executive Summary, July 2008.

[4] PDS 2010 Project Plan, March 2009.

2 Strategy

2.1 System Transition

System transition planning started with an effort of tradeoffs and impacts analysis. The table below summarizes the analysis.

	Data Providers	Data Users	Discipline Nodes	Engineering Node
Big Bang	Capabilities and services are delayed, however, greater functionality would be in place when they begin to deliver PDS4 data	Data users would get data in PDS4 format later, but could potentially have more tool support	Larger impact on the nodes if all nodes need to put services and capabilities in place first	Larger impact on the Engineering Node to put services and tools in place and ensure PDS-wide readiness to accept data for all new missions in PDS4 at once
Multi-phased	Capabilities can be put in place earlier and begin flow of PDS4 data and support	Data users would get data in PDS4 format earlier, but would possibly have less tool support	Less impact on the nodes if data is accepted in stages	Less impact on the Engineering Node and improved project performance since results can be realized earlier and PDS can deploy services and train over time

Opting for minimum impact and considering that current PDS being an operational system, the strategy is to employ multiple opportunities to deploy and transition from current PDS to PDS 2010, namely, multi-phased. Three system build releases have been identified. The release phasing details are described in the PDS 2010 Project Plan. The releases follow closely by development phases planned and mapped to components decomposed from the PDS 2010 system architecture. The transition periods align with the system releases. Complete deployment is planned in 2011 to accept PDS4 data. The acceptance of PDS4 data from new missions will be carefully staged.

2.2 Data Migration

PDS data migration from PDS3 to PDS4 must also be carefully considered. Migration planning also started with an effort of tradeoffs and impacts analysis. The tables below summarize the analysis.

	Resources	**Training**	**Usability**	**Efficiency**
No Migration	No impact on resources to convert data, however, PDS software will need to support PDS3 and PDS4	Users will need to be capable of working with PDS3 and PDS4 data	Limited support for working with PDS3 data in the future	Most cost-effective solution
On Demand Migration	Impact on conversion of a subset of critical data sets; PDS software will need to support PDS3 and PDS4	Users will need to be capable of working with PDS3 and PDS4 data	Limited support for working with PDS3 data in the future; critical data sets will be converted to PDS4 to improve usability	More costly than "No Migration", however, improves usability following a pragmatic approach
Full Migration	Substantial impact in converting data, redelivering to NSSDC, and developing supporting software	Users will ultimately need to only be familiar with PDS4	Usability would be improved since PDS data will be brought up to date	Substantial costs in migrating all data

	Data Providers	Data Users	Discipline Nodes	Engineering Node
No Migration	No impact	Users will need to be familiar with PDS3 and PDS4 data formats	Nodes will need to continue to provide support for PDS3 data	Engineering Node must continue to provide software support for PDS and PDS4
On Demand Migration	No impact	Users will need to be familiar with PDS3 and PDS4 data formats. However, critical data sets can be migrated to enhance usability.	Nodes will need to continue to provide limited support for PDS3 data; minor impact in migrating critical data sets	Engineering Node must continue to provide software support for PDS and PDS4
Full Migration	No impact	Data users will need to eventually only learn PDS4	Substantial impact in migrating and redelivering all data to PDS4	PDS3 tools and services can be retired once migration is complete

Opting for minimum impact and considering that current PDS being an operational system, the strategy is to execute migration of PDS3 data to PDS4 on demand, and on an as needed basis. Operational system will continue to support legacy pipeline to accept PDS3 data, allow users to access PDS3 data, and provide science services for PDS3 data. The planning and schedule for migration will exist within the Discipline Nodes.

3 Concept

3.1 System Transition

First, focus is placed on completion of initial PDS4 Standard (v4.0). We will work with new mission startups to define PDS4 products for future deliveries (e.g., 2012 and beyond). We will also work with International Planetary Data Alliance (IPDA) for adoption of the core when tools, infrastructure and standards in place.

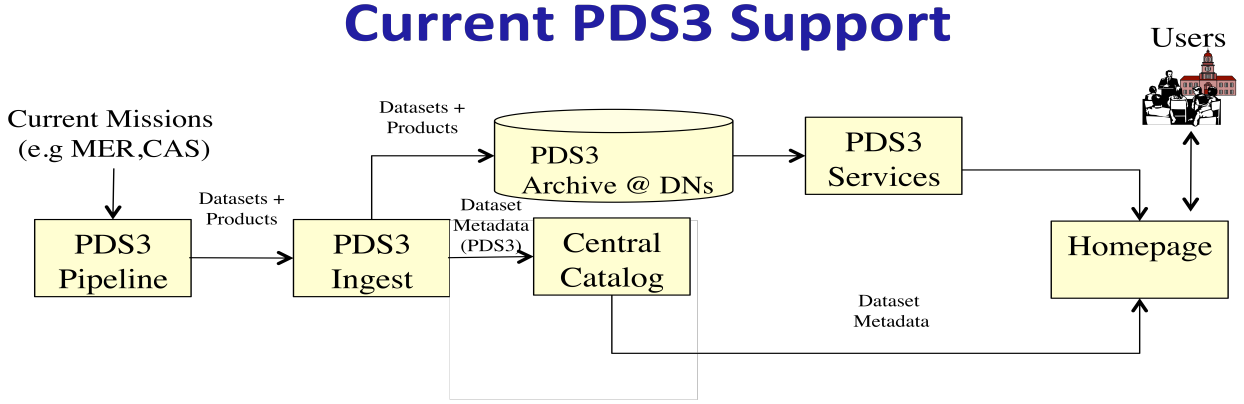
Validation of PDS4 will occur during first two system build phases. Version 4.0 data standards will be baselined with the first build of the system. This provides a standard to support development of PDS4 products for new mission startups. It will also support and validate PDS 2010 system development. This approach allows for phased transition to PDS4 over time.

As part of the PDS 2010 system transition, PDS Central Catalog will be replaced with a registry system that will support both PDS3 and PDS4 data collection and product registration. Tools will be provided to support PDS3 and PDS4 standards. Central PDS homepage will link to both PDS3 and PDS4 resources, as they are available.

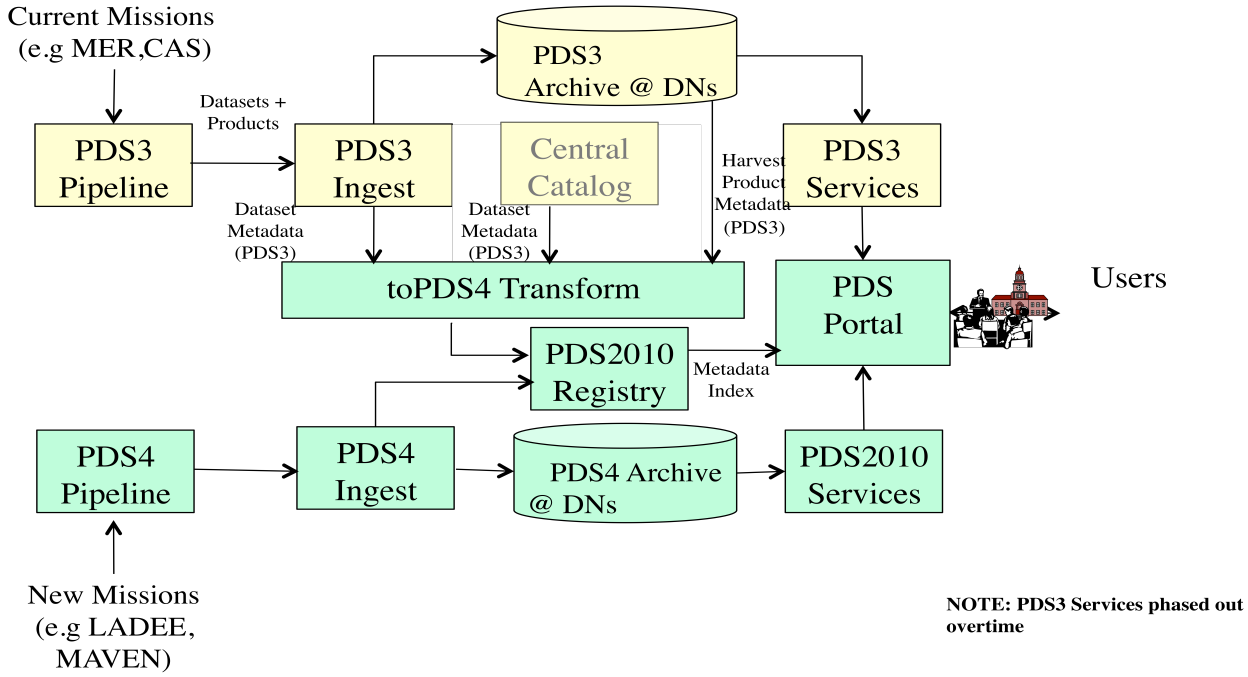
It is utmost important that existing PDS3 pipelines will be supported as we transition to new system and beyond. We will make sure that existing PDS3 pipelines will remain supported during life of mission. We will support ingestion and distribution of PDS4 data when ready. Every effort is made to ensure the new system will serve data from PDS3 and PDS4 repository. Missions and IPDA partners can transition when they want to. Overall plan is that existing PDS3 services will remain while new PDS4 services will be added. Each node will execute their own transition timeline in term of providing new PDS4 services.

The following diagram depicts current PDS3 support in place today, follows by the diagram that depicts the transition concept:

Current PDS3 Support



Transition to PDS4 Support



NOTE: PDS3 Services phased out overtime

3.2 Data Migration

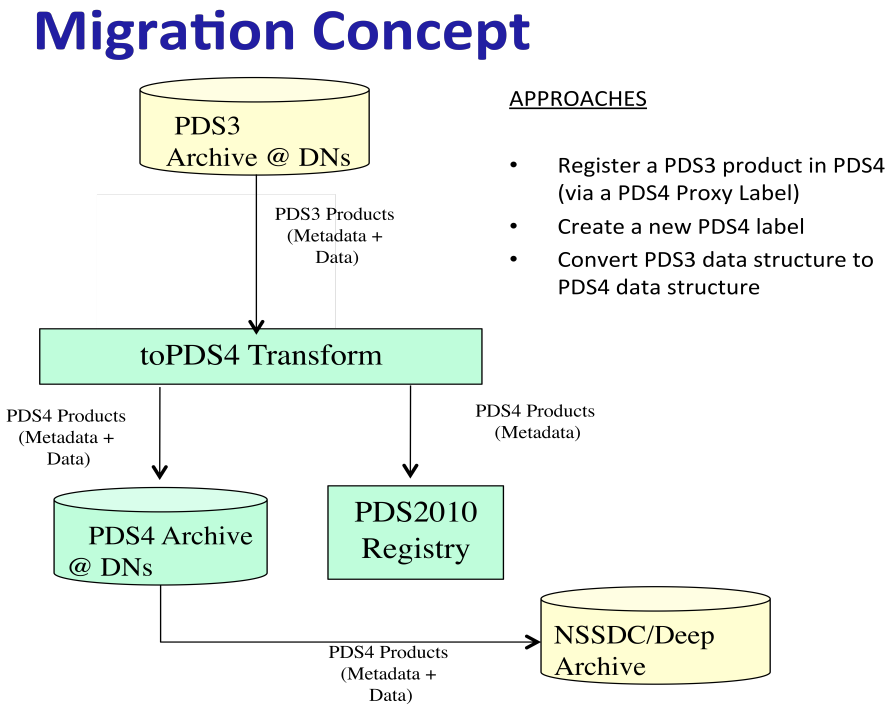
Migration of PDS3 data to PDS4 is to occur on demand, on an as needed basis. Data migration entails on-going planning of what existing PDS3 data and when they best be migrated to PDS4, as well as staging when to accept PDS4 data from what missions. In addition, a well-defined data migration process will be established to support the migration of exiting PDS3 data to PDS4. The actual data migration tasks will be planned and executed by the Nodes.

A Data Migration Process will be established to facilitate the migration. Depending on the level of complicity, data will be converted by one or more of the following methods:

1. Convert PDS3 label to PDS4 label and register the PDS4 label in the PDS 2010 registry without changing the data object of the product
2. Convert PDS3 label to PDS4 label and register the PDS4 label in the PDS 2010 registry plus convert the PDS3 data object to PDS4 data object

Migrated data will be sent to the National Space Science Data Center (NSSDC) for deep archive.

The following depicts the migration concept:



NOTE: Deliveries to the Deep archive will be performed if actual label and/or data has been transformed.

4 Stakeholders

The transition to PDS 2010 system affects many different organizations involved with PDS. The following table lists those groups, summarizes their roles, and involvement that the transition requires of them.

Stakeholder Group	Role	Transition Involvement
Engineering Node (EN)	PDS 2010 system and PDS4 development and deployment	The EN is responsible for deploying PDS 2010 system, helping data providers to use the new standards and tools, and support Discipline Nodes integration and migration.
PDS Discipline Nodes (Nodes, DN)	Support PDS 2010 system and PDS4 development and deployment	The PDS Discipline Nodes (DN) are responsible for helping data providers to use the new standards and tools, supporting integration of DN system with the new system, data migration, and development of discipline specific science services.
Data Providers	PDS3 and PDS4 products generation and delivery	The PDS4 data providers are responsible for construction of PDS4 mission pipeline, working with DNs to generate and deliver PDS4 products.
Data Consumers	Use PDS3 and PDS4 data	The data consumers use new system to search and access both PDS3 and PDS4 data.
International Planetary Data Alliance (IPDA)	Support international archives	International community develops PDS4 derived interoperable planetary science archives.

5 Activities

There will be three transition periods. They align with the PDS 2010 project build releases. The release phasing details are described in the PDS 2010 Project Plan. The tasks and activities required to support each transition period are listed below.

5.1 Transition Period 1

The first transition period begins when we are ready to deploy PDS 2010 prototype build 1 release. The period will end when the operational PDS 2010 build 2 release is ready to be deployed.

Engineering Node tasks

Prototype build 1 – Install PDS 2010 prototype build 1 which includes the PDS 2010 Ingest subsystem (Harvest, Registry, Security), Report service, PDS4 Validation tools, and Administrative portal, and documentation (release description, design, installation, etc).

PDS4 Standards - Provide a baseline version of PDS4 Model Specification, Standards Reference, Data Dictionary, Schemas, Data Provider's Handbook, as well as Tutorials; participate in International Planetary Data Alliance (IPDA) projects to ensure that IPDA model is compliant to PDS4 core standards.

Training - Provide PDS4 standards training sessions to PDS4 data provides and Nodes

Development – Develop build 2 system (product generation tools, integrated PDS3 and PDS4 product search and distribution, etc); extend the PDS4 standards to support and integrate the Node specific data models.

Maintenance – Continue to support PDS3 pipelines. Maintain PDS3 tools, PDS3 archive, Central Catalog and Home Page. Maintain PDS3 standards.

Data Engineering – Continue to support PDS3 validation and ingest into the Central Catalog; participate in PDS4 standards training; exercise PDS4 tools and provide feedback; support reconciliation of Central Catalog and the prototype PDS 2010 Registry.

Discipline Node (DN or Node) tasks

Prototype build 1 – Download and install Harvest client (it will crawl the Node's PDS3 repository and register the products metadata at the prototype remote Registry at EN); register prototype PDS4 products using remote Registry at EN; download prototype validation tools. For selected Nodes, install prototype local Registry service.

PDS4 Standards – Begin to develop Node specific PDS4 data model.

Training – Participate in PDS4 standards training sessions; provide feedback of the standards and model.

Development – Begin to plan and develop capabilities needed to integrate with PDS4 data access and search, including migration of existing catalog to local PDS 2010 registry, and updating Node specific services (e.g. data search web pages). Also, start to plan and develop PDS4 archive.

Maintenance – Continue to support PDS3 pipelines. Maintain DN system and PDS3 archive.

Data Engineering – Continue to support PDS3 validation, ingest into the Node PDS3 archive and submit for Central Catalog ingestion; exercise PDS4 validation tools and provide feedback to EN; create prototype PDS4 products and work with data providers; provide feedback to EN; begin to develop data migration plan.

Data Provider tasks

PDS3 products provider – Continue to use PDS3 tools to create and validate PDS3 products and submit to Nodes.

Potential PDS4 products provider - Participate in PDS4 standards training sessions; provide feedback of the standards and model, work with Nodes to create prototype PDS4 products; exercise PDS4 validation tools and provide input to PDS.

Data Consumer tasks

Continue to use existing PDS3 tools and science services; continue to access and search PDS3 products using existing Home Page.

International Planetary Data Alliance (IPDA) tasks

Start development of IPDA core model that aligns and will be compliant with PDS4.

5.2 Transition Period 2

The second transition period will start when we ready to deploy our first operational PDS 2010 system (build 2 release), and it will end when we are ready to deploy build 3 release.

Engineering Node tasks

Operational (build 2) release – Deploy PDS 2010 operational system (complete data management infrastructure: ingest, search, distribution) and PDS4 tools suite (validation and generation of PDS4 products), and documentation.

PDS4 Standards - Provide a release copy of PDS4 Model Specification, Standards Reference, Data Dictionary, Schemas, Data Provider's Handbook, as well as Tutorials;

continue to participate in International Planetary Data Alliance (IPDA) projects to support development of IPDA model to ensure PDS4 compliance.

Training - Provide PDS4 standards training sessions to additional PDS4 data providers, Nodes and users.

Development – Develop build 3 (new PDS web site and common PDS4 science services).

Maintenance – Continue to support PDS3 pipelines. Maintain PDS3 and PDS4 tools, PDS3 archive, Home Page and PDS 2010 system; decommission Central Catalog. Maintain PDS3 and PDS4 standards.

Data Engineering – Continue to support PDS3 data providers; exercise PDS4 operational tools suite and provide feedback; support reconciliation of Node repositories and the PDS 2010 Registry; ready to support PDS4 data providers.

Discipline Node (DN or Node) tasks

Operational (build 2) release – Install PDS 2010 Registry; download PDS4 tools suite; ready to support ingest of PDS4 products.

PDS4 Standards – Continue to provide Standards feedback to EN; support integration of Node specific data model with PDS4 core model.

Development – Continue to develop capabilities needed to integrate with PDS4 data access and search; continue to develop PDS4 archive; start development of Node specific PDS4 science services.

Maintenance – Continue to support PDS3 pipelines. Maintain DN system and archive(s).

Data Engineering – Continue to support PDS3 validation, ingest into the Node PDS3 archive; work with PDS4 data providers; exercise PDS4 tools suite (validation and generation) and provide feedback to EN; ready to support PDS4 products ingest; continue to plan and execute data migration.

Data Provider tasks

PDS3 products provider – Continue to use PDS3 tools to create and validate PDS3 products and submit to Nodes.

Potential PDS4 products provider – Continue to participate in PDS4 standards training sessions; provide feedback of released standards and model, ready to work with Nodes to create PDS4 products; exercise PDS4 tools suite and provide input to PDS.

Data Consumer tasks

Continue to use existing PDS3 tools, search and science services; use available PD3/PDS4 integrated search and access capabilities to access and search PDS3 and PDS4 products; use PDS4 tools; provide input to PDS.

International Planetary Data Alliance (IPDA) tasks

Continue development of IPDA core model that aligns and will be compliant with PDS4; start to plan and develop IPDA archives.

5.3 Transition Period 3

The final transition period starts when we ready to deploy our operational PDS 2010 system's second major release (build 3).

Engineering Node tasks

Operational (build 3) release – Deploy enhanced PDS 2010 system and tools, PDS 2010 new website, and selected PDS common science services (e.g. output format transformation).

PDS4 Standards - Continue to participate in International Planetary Data Alliance (IPDA) projects to support development of IPDA model to ensure PDS4 compliance.

Maintenance – Maintain PDS3 and PDS4 tools, PDS3 archive, Home Page and PDS 2010 system. Maintain PDS3 and PDS4 standards.

Data Engineering – Continue to support PDS3 and PDS4 data providers; exercise enhanced PDS4 tools suite and provide feedback; continue to support reconciliation of Node repositories and registries and the PDS 2010 central registry at EN.

Discipline Node (DN or Node) tasks

Operational (build 3) release – Install build 3 PDS 2010 Registry and download tools; exercise and provide feedback to EN.

PDS4 Standards – Continue to enhance Node data model, support integration and provide feedback to EN.

Development – Continue to develop capabilities needed to integrate with PDS4 data access and search; continue development of Node specific PDS4 science services; continue to plan and develop PDS4 archive.

Maintenance – Continue to support PDS3 and PDS4 pipelines. Maintain DN system and archive(s).

Data Engineering – Continue to work with data providers; support PDS3 and PDS4 validation, ingest into the Node PDS3/PDS4 archives; continue to plan and execute data migration.

Data Provider tasks

PDS3 products provider – Continue to use PDS3 tools to create and validate PDS3 products and submit to Nodes.

Potential PDS4 products provider – Continue to participate in PDS4 standards training sessions; provide feedback of released standards and model, work with Nodes to create PDS4 products; exercise PDS4 tools suite and provide input.

Data Consumer tasks

Continue to use existing PDS3 tools, search and science services; use new PDS 2010 website to access and search PDS3 and PDS4 products; use PDS4 tools; provide input to PDS.

International Planetary Data Alliance (IPDA) tasks

Continue development of IPDA core model that aligns and will be compliant with PDS4; start to plan and to develop IPDA archives.

Appendix A: Acronyms

CAS – Cassini Mission
DN – PDS Discipline or Data Node
EN – PDS Engineering Node
IPDA – International Planetary Data Alliance
MER – Mars Exploration Rover missions
NASA – National Aeronautics and Space Administration
Node – PDS Discipline Node
NSSDC - National Space Science Data Center
PDS – Planetary Data System
PDS3 – Version 3.8 of the PDS Data Standards
PDS4 – Version 4.0 of the PDS Data Standards
PDS 2010 – PDS 2010 Project
PDS MC – PDS Management Council