Planetary Data System Strategic Roadmap 2006 - 2016

PDS Management Council

Planetary Data System Mission

To facilitate achievement of NASA's planetary science goals by efficiently archiving and making accessible digital data produced by or relevant to NASA's planetary missions, research programs, and data analysis programs.

Feb 2006

http://pds.nasa.gov/

Topics Covered in This Presentation

- NASA's Vision for the Planetary Data System (PDS)
- Characteristics of The PDS
- Missions in Progress
- Current Challenges
- Ongoing Challenges
- Requirements
- Functions
- Implementation
- PDS Realities
- The PDS 5-year Goals
- Milestones to Implement
- PDS in 5 Years

NASA's Vision for the PDS

To gather and preserve the data obtained from exploration of the Solar System by the U.S. and other nations

To facilitate new and exciting discoveries by providing access to and ensuring usability of those data to the worldwide community

To inspire the public through availability and distribution of the body of knowledge reflected in the PDS data collection

To support the NASA vision for human and robotic exploration of the Solar System by providing an online scientific data archive of the data resources captured from NASA's missions

Characteristics of the PDS Data Archive

The PDS archives and makes available space-borne, ground-based, and laboratory experiment data from over 50 years of NASA-based exploration of comets, asteroids, moons, and planets.

The archives include data products derived from a very wide range of measurements, e.g., imaging experiments, gravity and magnetic field and plasma measurements, altimetry data, and various spectroscopic observations.

Planetary missions frequently have short or intermittent observing phases that result in limited accuracy of calibration and incomplete understanding of instrument stability and characteristics.

Many of the data sets are unique in that the observations cannot be duplicated.

The wide range of archival products and associated disciplines are fundamental reasons why PDS is organized as a federation serving various science communities.

The future is likely to be as diverse, unpredictable, and challenging as the past - requiring innovative management and state-of-the-art technology within a stable yet robust archiving framework.

Missions in Progress

Mission	FY06				FY07				FY08				FY09				FY10				FY11				FY12			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Cassini-Huygens	E																											
Dawn (Discovery 9)					E																							
Deep Impact (Discovery 8)																												
Deep Space 1	E																											
Discovery AO-2006 (Discovery 12)					A/B												C / D											E
GALEX (SMEX 7)	E																											
Genesis (Discovery 5)	E																											
Hayabusa-U.S. (MUSES-C)	E																											
Juno (New Frontiers 1)									C / D												E							
Kepler (Discovery 10)	C / D								E																			
Lunar Lander Project	A/B								C / D												E							
Lunar Reconnaissance Orbiter	C / D												E															
Mars Exploration Rovers	E																											
Mars Express																												
Mars Global Surveyor	E				proposed				proposed																			
Mars Science Laboratory									C / D								E											
Mars Odyssey	E																											
Mars Reconnaissance Orbiter	E																											
Mars Scout 2 (launch 2011)						A/B											C / D											E
Mars Sample Return (2013)									A/B												C / D							
MESSENGER	C / D				E																							
Moon Mineralogy Mapper	C / D								E																			
Muses C	C/D																											
NEAR	E																											
New Frontiers 2									A/B												C / D							
New Horizons									C / D																			
Phoenix (Mars Scout 1)	C / D								E																			
Rosetta	E																											
Spitzer Space Telescope (SIRTF)	E																											
Stardust (Discovery 4)	E																											
Ulysses	E																											
Venus Express (limited US involvement)	E																											
Voyager	E																											

Note: Letters and colors correspond to NASA mission phases

Feb 2006

Mission and Data Provider Challenges

Data volumes will continue to increase

- Mars Reconnaissance Orbiter will produce ~100 TB
- Mars data volume is likely to be surpassed by that returned from the Moon in the next 10 years.

Missions, instruments and data are becoming more complex

- Rovers generate complex sets of products
- Flagship Missions strive to satisfy many science goals leading to a variety of instruments, complex structure of data returned, and extended missions with reduced budgets for archive production

Smaller or more focused missions present challenges

- The budget is limited and funding profile differs from that used in initial design & planning
- PDS must supply tools to help with data archiving task, provide a Common Data Model, and participate as early as possible
- Mission phases are compressed and require rapid PDS response

User and Customer Challenges

User desires and expectations are becoming more complex and demanding with time as data and mission complexity increase

- Users want and need derived products
- Users expect rapid access, flexible search capability and refined catalogues
- Users want readily accessible data visualization, analysis, and cross-correlation tools
- Experienced users frequently want interfaces with familiar data manipulation packages
- Inexperienced users are "Google" oriented but PDS data is not always adaptable to the expected indexing .

Interdisciplinary Science requires parametric, geographic or spatial searches across missions, instruments and data sets

Lack of reduced data records, combined with less than adequate search and retrieval capabilities associated with the current implementation of PDS sometimes frustrate users who wish to find and download highly processed data for further analyses

Challenges Associated with International Collaboration

Planetary exploration has become increasingly international in scope.

- International missions providing the data
- International scientists using the data

The science community expects seamless access to integrated planetary data archives across international boundaries which requires

- Plans and agreements for sharing data
- International data standards
- International protocols for sharing data

Currently, there are no international data standards for planetary science

- PDS has made progress in having its standards become internationally accepted, e.g., ESA's Huygens, Mars and Venus Express data sets will be compliant with PDS.
- The Planetary Science Archive (PSA) and PDS are working on establishing such a standard based on PDS

Coordination and development and application of standards for archive production across international partners substantially increases workload due to increasing number of interfaces and standards training

ITAR issues that may preclude fully open discussion among international collaborators

Operational Challenges

Users require that PDS supports heterogeneity for both technology and science

- They need software tools that can run on a wide variety of platforms and adapt for different uses (E.g., different mission and instrument teams)
- They need to be able to use PDS for discipline-specific science both in archiving and searching and retrieving data

Users expect to be able to search all of the PDS data holdings within the system and perform correlated searches for products across nodes

Users expect online access and distribution of all the data holdings

- Users need the ability to download data products and volumes of increasing sizes
- PDS needs the ability to move large volumes between institutions

Upcoming missions require that PDS be able to scale critical functions

- The storage infrastructure must scale to support the increase of volume
- The architecture must scale to support the complexity of missions flying many different types of instruments
- The operations of PDS must scale to support the increasing number of missions that are occurring concurrently

Aging and obsolete technology must be refreshed to ensure long term usability of the data and the system

Data in the archive must be periodically validated to ensure long term preservation

Ongoing Challenges

Within a "flat" funding profile, PDS must:

- Interface with new missions (typically more than 20 at one time in various development phases)
- Ingest data from increasingly complex instruments
- Add non-mission data from individuals, labs and observatories
- Incorporate new storage/distribution technologies
- Educate data providers
- Provide useful links to other, related data systems
- Respond to an increasing spectrum of demands from an ever-growing community of users

Requirements

PDS will provide expertise to guide and assist missions, programs, and individuals to organize and document digital data that can be used to support NASA's goals in planetary science and Solar System exploration

PDS will collect suitably organized and well-documented data into archives that are peer reviewed and maintained by members of the scientific community.

PDS will make these data accessible to users seeking to achieve NASA's goals for exploration and science

PDS will ensure the long-term preservation of the data and maintain their usability

See backup slides 1-3 for details

PDS Functions

PDS serves the scientific community by assuring the availability of high quality and scientifically useful data products

To accomplish these goals, PDS

- **works with Data Providers to Prepare Archival Quality Data Products**
- **provides Access to Data from NASA and International Missions**
- **delivers Data to the Scientific Community**
- **establishes a Common Data Model and Data Dictionary for Planetary Data**
- **sets Archival Standards**
- **preserves the Data**
- **assists Scientists in Accessing and Using Planetary Data**
- **is responsive to a diverse community of users**
- **facilitates Education and Public Outreach**

Implementation - Operation

PDS is a close federation Nodes with both Science and Support functions

- Science functions are organized by discipline and include
  - data ingestion
  - data distribution
  - interfacing with data suppliers and users to ensure that
    - maximum science value is captured within the archive
    - the archive is of greatest utility to both immediate and long-term science users
      - Immediate users, by their use of the system, help PDS understand if the services and data sets are of optimal use to the community. These users have the benefit of an active instrument team to whom comments and replies can be passed, if needed, allowing the archive to be modified.
      - Long-term users need final, stand-alone archives because the instrument experts may no longer be available.

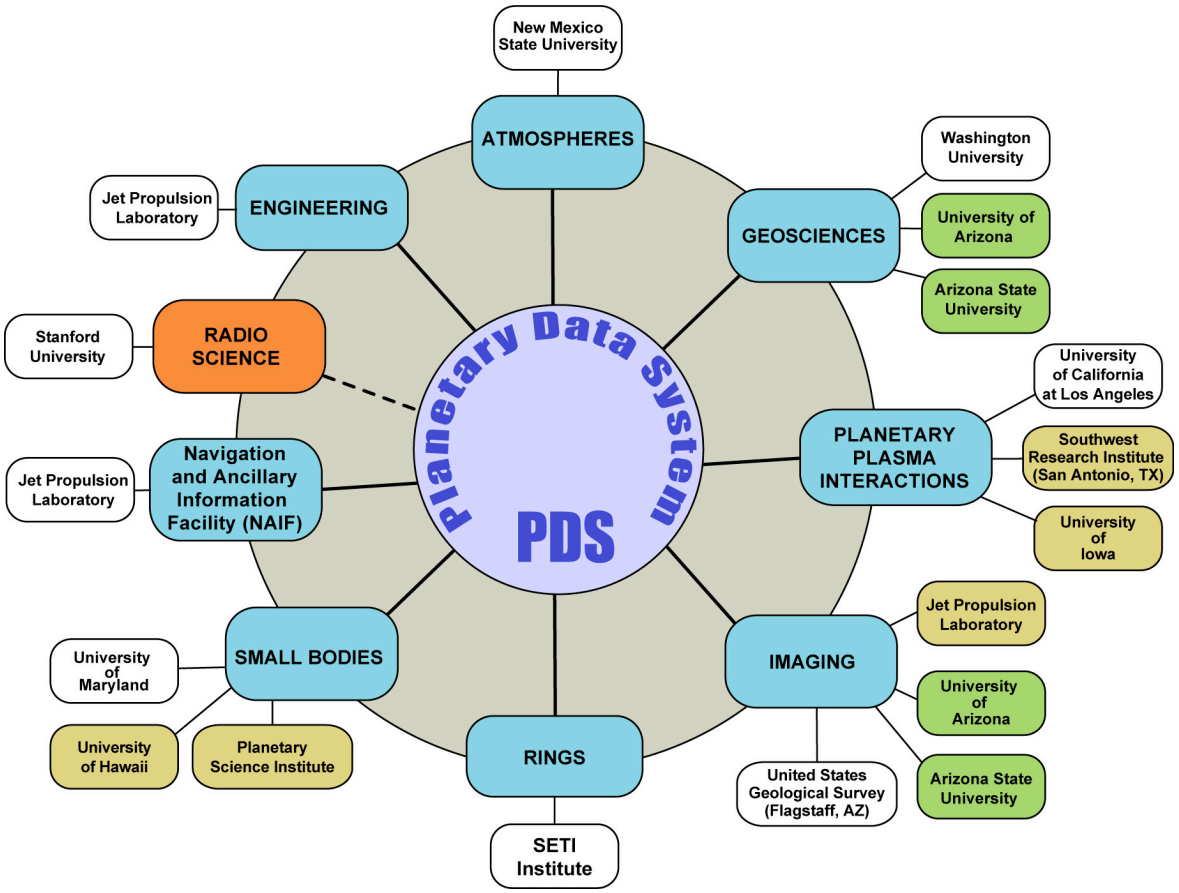
Implementation - Operation (cont.)

- **Support functions include basic development and cross-discipline support such as**
  - **common tools, libraries, procedures, and standards for data preparation, submission, and management**
  - **common tools for data manipulation**
  - **an infrastructure that facilitates**
    - **easy navigation within and access to holdings throughout the federation**
    - **simple system-wide maintenance and upgrades**
- **The Engineering Node focuses primarily on Support; the other Nodes focus primarily on Science, but all have at least test bed and review tasks**

Implementation - The PDS System

NODES/SUBNODES/DATA NODES

Function



Feb 2006

Implementation - Education and Public Outreach

All tools and holdings of the PDS are available to everyone, but the primary customers are scientists and their upper division and graduate students and associates.

PDS Nodes maintain tools that support advanced class work and research projects.

PDS directs users to other valuable related resources

PDS evaluates and accepts requests for product generation and delivery for EPO Projects

Backup sides 4-5

Implementation - Assessment and Evaluation

External community assessment is often based on user expectations that do not match requirements and charter

- **Well defined criteria for evaluation must be implemented**
- **Defensible metrics need to be presented**

The PDS will maintain an ongoing internal evaluation addressing effectiveness of all PDS functions. Data to support the evaluation process will be assimilated from:

- **Regular consultation with individual Node Advisory Groups**
- **Periodic surveys of the data provider and data user communities**
- **Unsolicited user feedback**
- **Storage, access, and retrieval statistics**

PDS Realities

Inadequate Mission budgets for archiving have caused problems for PDS since its inception

- Data archiving in general has been relatively low-priority for many missions. Increased costs and unexpected mission problems have often pushed archiving further down the priority list.
- Instrument teams have been limited by time and funds in their ability to generate, validate, and deliver to the PDS derived products of use to the science community
- Most missions did not plan adequate funding for development and delivery of derived products.
- Data analysis programs have not been funded to generate, validate, and deliver-compliant derived products to the archive

The PDS 5-Year Goals

Significantly improved user services based in part on periodic user feedback, data searches correlated across multiple data sets, new data visualization tools and enhanced delivery capabilities

Each NASA planetary mission plan end-to-end archive generation and achieve on-time delivery of documented and validated PDS-compliant archives

Generation, validation, and delivery of PDS-compliant derived products by NASA mission instrument teams and PI's from data analysis programs as a standard practice

Establish and maintain coordination with domestic and international counterparts to assure interoperability making available data from all planetary missions (i.e., NASA, ESA, and other agencies)

Milestones to implement 5-Year Goals

Data Access and Distribution

- PDS will support a reduced set of file structures/formats for science data, with the selection based on community input.
- All Discipline Node catalogs will provide search capabilities across all their holdings at a fine level of granularity, including geometric constraints.
- Basic data visualization tools will be integrated into the search process, so that most data products can be evaluated quickly by the user before selection for delivery.
- Automatic translation tools will allow users to receive data in whatever suitable format they prefer.
- Users will have the option of receiving data requests in the form of fully PDS-compliant custom volumes generated on-the-fly.

Milestones to implement 5-Year Goals (cont.)

Team Interfaces & Archiving

- Mission archiving with PDS will be routine
  - Well established milestones that evolve into place based on PDS and community experience and inputs
  - Discipline Nodes will provide tools to the teams for generating the necessary metadata relevant to their discipline
  - A complete set of automated tools will be available to validate the metadata against PDS standards

- AOs need to require all instrument/mission teams to
  - deliver fully calibrated, geometrically corrected data throughout the mission and to resubmit the data when they are improved
  - collaborate with the PDS on all data calibration/processing/retrieval systems designed for their team members. Data systems used by team members must seamlessly transition over to the PDS as the team's data sets become publicly available.

Milestones to implement 5-Year Goals (cont.)

Technical Infrastructure

- PDS architecture will be entirely distributed. Facilities at each Discipline Node will be capable of handling cross-discipline queries
- All data holdings will reside at a minimum of two different, geographically separated locations. Popular data sets will be continuously on line at two or more locations
- All PDS software will be written in a manner that minimizes, and preferably eliminates, hardware and OS dependencies
- PDS will upgrade its network connections regularly to provide fast delivery of data to users
- PDS will collaborate with other NASA entities to better assess and utilize existing and emerging storage technologies

To achieve these goals PDS will continue and enhance its proactive role in working with NASA Headquarters and Mission management and instrument teams to identify and maintain necessary processes, schedules, data models, standards and assistance.

The PDS in 5 Years

PDS is committed to a major, structured modernization of its data archiving & distribution

The framework will be PDS-4; key aspects of which will include:

- **Inputs from the mission & science user communities**
- **A simplified set of data formats**
- **Metadata requirements to support modern search technology**
- **Highly automated validation & ingestion**
- **On-the-fly conversions from stable, streamlined archival formats to suitable, user requested formats**
- **Improved interoperability with other domestic and international space data systems**

The pace and fidelity of this development will depend on funding levels and headquarters support through the inclusion of specific requirements in future AOs

Backup Materials

PDS Level 1 & Level 2 Requirements

PDS will provide expertise to guide and assist missions, programs, and individuals to organize and document digital data supporting NASA's goals in planetary science and Solar System exploration

- **1.1 Single Point of Contact:** PDS will provide a single point of contact to each mission, program, agency or individual (i.e., data providers) wishing to submit archival data
- **1.2 Expert Help:** PDS will provide expert help in designing archival data sets
- **1.3 Plans and Documents:** PDS will assist data providers in developing archive plans, interface control documents, validation procedures, and delivery schedules
- **1.4 Archiving Standards:** PDS will develop, publish, and maintain archiving standards for planetary science data
- **1.5 Archiving Tools:** PDS will develop, distribute, and maintain tools to assist data producers in assembling, validating, and submitting archival products

Will be updated as new versions are developed

PDS Level 1 & Level 2 Requirements (cont.)

PDS will collect suitably organized and well-documented data into archives that are peer reviewed and maintained by members of the scientific community.

2.1 Receive: PDS will develop, publish, and implement procedures for receiving, acknowledging and tracking data submissions

2.2 Validation: PDS will develop, publish, and implement procedures for validating data submissions to ensure compliance with standards

2.3 Peer Review: PDS will develop, publish, and implement procedures for conducting peer reviews of all data submissions to help ensure completeness, accuracy and usability of content

2.4 Acceptance: PDS will develop, publish, and implement procedures for accepting or rejecting archival data

2.5 Catalog: PDS will maintain a catalog of accepted archival data sets

2.6 Storage: PDS will develop and maintain appropriate storage for its archive

PDS Level 1 & Level 2 Requirements (cont.)

PDS will make these data accessible to users seeking to achieve NASA's goals for exploration and science

3.1 Search: PDS will allow and support searches of its archival holdings

3.2 Retrieval: PDS will facilitate transfers of its data to users

PDS will ensure the long-term preservation of the data and maintain their usability

4.1 Long-Term Preservation: PDS will determine requirements for and ensure long-term preservation of the data

4.2 Long-Term Usability : PDS will establish long-term usability requirements and implement procedures for meeting them

Education and Public Outreach - Tools

Current PDS tools support advanced class work and research projects.

- MER Analyst's Notebook (http://anserver1.eprsl.wustl.edu/) with data collection, viewable images and analysis notes by and for scientists
- Encyclopedia of Planetary terms (http://atmos.nmsu.edu/)
- Planet viewer, moon trackers and ephemeris generators (http://pds-rings.seti.org/tools/) to use for planning observations, writing proposals and preparing for public nights at educational facilities.
- NASAView (http://pds.jpl.nasa.gov/tools/software_download.cfm) for viewing imaging data
- Planetary Image Atlas (http://pds-imaging.jpl.nasa.gov/Atlas/) for assessing imaging data
- Map-a-Planet (http://pdsmaps.wr.usgs.gov/maps.html/) for selecting specific area maps
- PDS documents page (http://pds.jpl.nasa.gov/documents/) for tutorials on PDS functions
- Comet modeling tools http://pds-smallbodies.astro.umd.edu/comettools/index.html

Education and Public Outreach- Related Sites

PDS directs users to other valuable NASA resources

- The Planetary Photojournal (http://photojournal.jpl.nasa.gov/) with accessible images and accurate captions provides a timely resource for students and educators
- The NASA Astrophysical Data System (http://adswww.harvard.edu/) that provides a detailed abstract search
- Imaging and Geosciences nodes provide primary mission interfaces for NASA Regional Planetary Information Facilities (RPIFs) (http://www.lpi.usra.edu/library/RPIF/)
- ISIS web site (http://isis.astrogeology.usgs.gov/) widely used USGS developed software for processing image arrays and cubes
- NASA/USGS Planetary GIS Web Server (http://webgis.wr.usgs.gov/)
- Gazetteer of Planetary Nomenclature (http://planetarynames.wr.usgs.gov/) generated by the International Astronomical Union
- Orbital data for small bodies at the IAU's Minor Planet Center at CFA (largely NASA funded). http://cfa-www.harvard.edu/cfa/ps/mpc
- JPL's Horizons system for orbital data and ephemeredes and to JPL's orbit visualization tool. http://ssd.jpl.nasa.gov/
- A variety of sites for Near-Earth Asteroids, including sites at JPL. http://neo.jpl.nasa.gov/ . http://neat.jpl.nasa.gov/
- A tool for finding comet data in the HST data base http://pdssbn.astro.umd.edu/sbnhtml/comets/HST/