NASA's Planetary Data System Status and Future Plans

Tom Morgan and Maria Banks
For the PDS team
Report to the Planetary Advisory Council

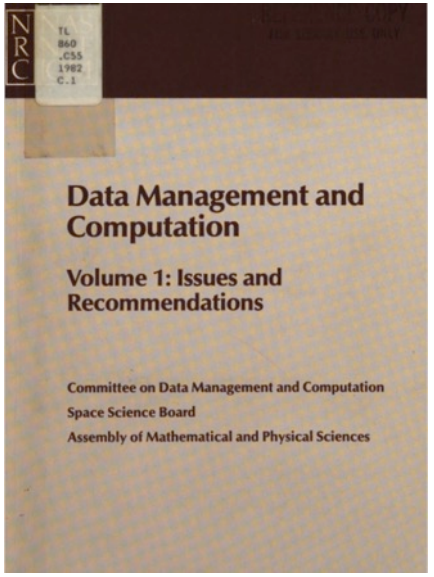
February 22, 2018

I. Planetary Data System (PDS): Overview (Tom)
II. NASA Planetary Mission Data Status (Tom)
III. PDS Roadmap Study: Overview and findings (Maria)
IV. Post Roadmap activities (Maria)
V. Conclusions and Future Plans (Tom)

Overview - Founding the PDS

Planetary Data System (PDS):

- Established by NASA in 1989
- A solution to concerns that the data being returned by scientific satellites was in danger of being lost



Challenges:

- data storage media
- adequacy of documentation
- lack of data availability outside of the implementing science team
- lack of any consistent standards for long-term data archiving

The National Academy of Sciences chartered the Committee on Data Management and Computation (CODMAC) in 1982:

- CODMAC issued three reports detailing means to address what were identified as serious problems in the way that NASA was managing its planetary data holdings.
- Central among the recommendations was to have **a scientifically guided distributed data system**, **adequately funded both to archive data and distribute it to researchers in a timely fashion.**

Overview - Founding the PDS continued

Following the National Academy's recommendations, the PDS uses:

I. **Discipline oriented nodes**, competed periodically (most recently in 2015)
II. **Two support nodes**, the Engineering node and the Navigation and Ancillary information node, both at JPL
III. **Small project office at GSFC** to manage funds, and to coordinate activities

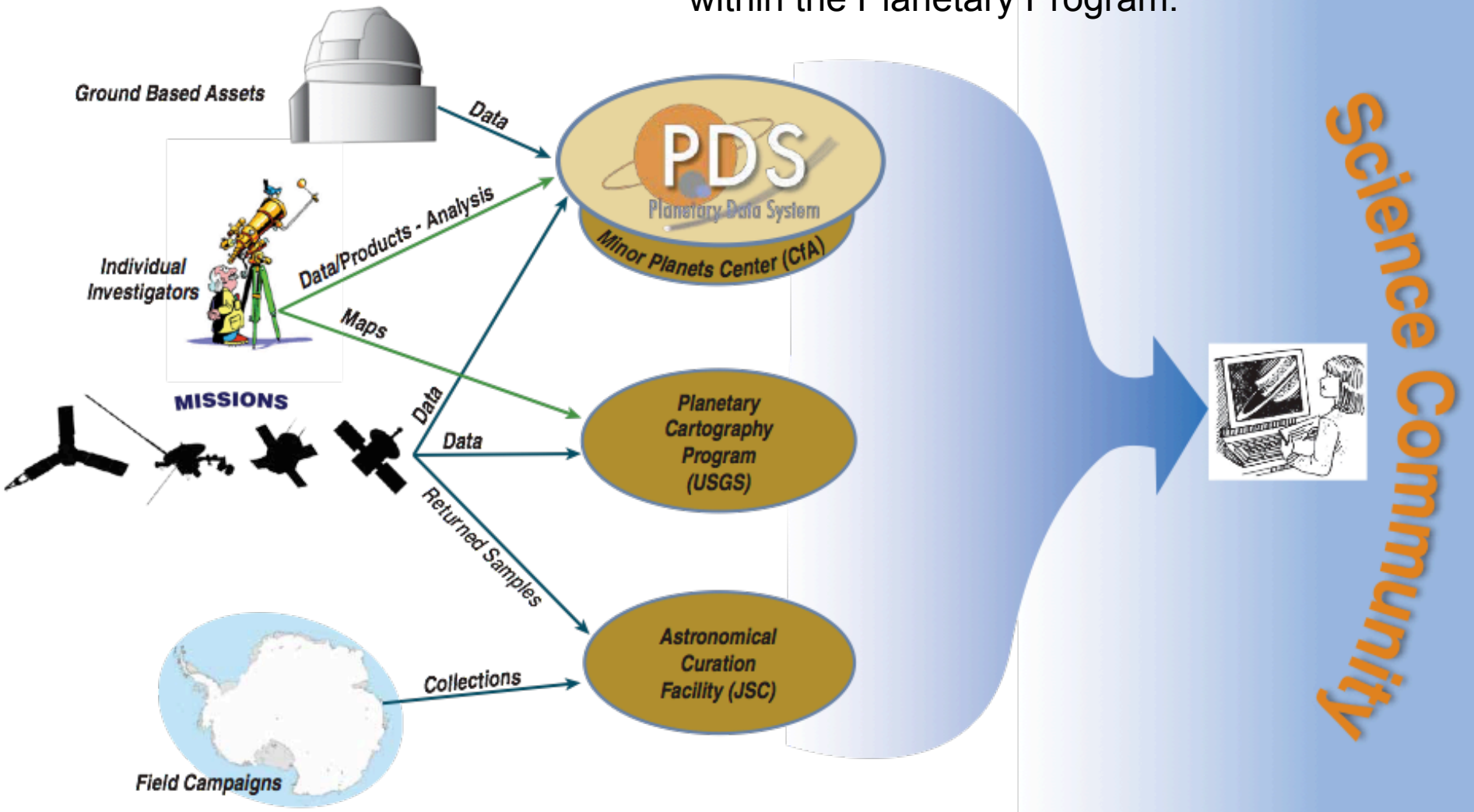
The PDS is the archive for NASA's planetary missions, but it is only one of 16 formal NASA Archives, plus 22 Guest Observer Facilities and Science Centers, supporting all of the science missions within Science Mission Directorate within NASA. Many of these archives, like PDS, include specialized organizations within the archive to accomplish that archive's requirements.

- The PDS maintains for current and future use a diversity of unique data sets acquired as national assets from across the Solar System after the expenditures of multiples of billions of dollars of U.S. Government investment.

- PDS works with every NASA planetary mission with the goal of obtaining a complete archive of data from that mission.

Overview - What is the Planetary Data Environment Like?

The PDS is one of three complementary archival elements within the Planetary Program.



Overview - Archiving
Focus of the PDS

Purpose: Collect, archive and make accessible the digital data and documentation produced from NASA's exploration of the solar system from the 1960s to the present.

The PDS actively manages the archive to maximize its usefulness, and it has become a basic resource for scientists around the world.

- All PDS-curated products are peer reviewed, well documented, and available online to scientists and to the public without charge.
- Online search capabilities are provided.
- The PDS is an actively accumulating data archive.

Overview - Archiving Focus continued

- The PDS uses **stable standards** for describing and storing data that are designed to enable future scientists who are unfamiliar with the original experiments to analyze the data.

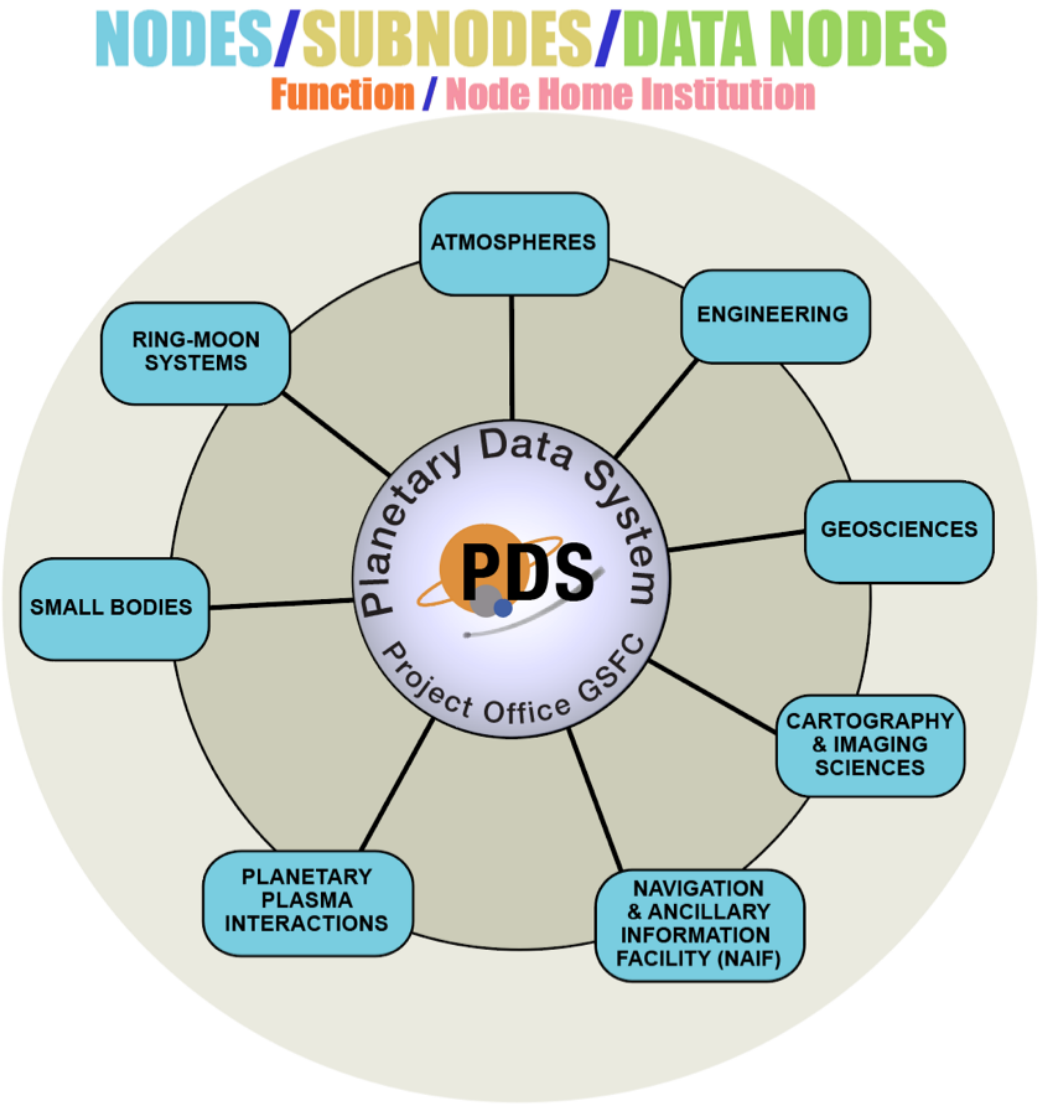
These standards address the data structure, description contents, media design, and a set of terms.

- **PDS Archiving Standard System (PDS4) is based on a well defined planetary science information model.** The system is extensible, and there is a Change Control Board to ensure that changes are consistent with the information Model.

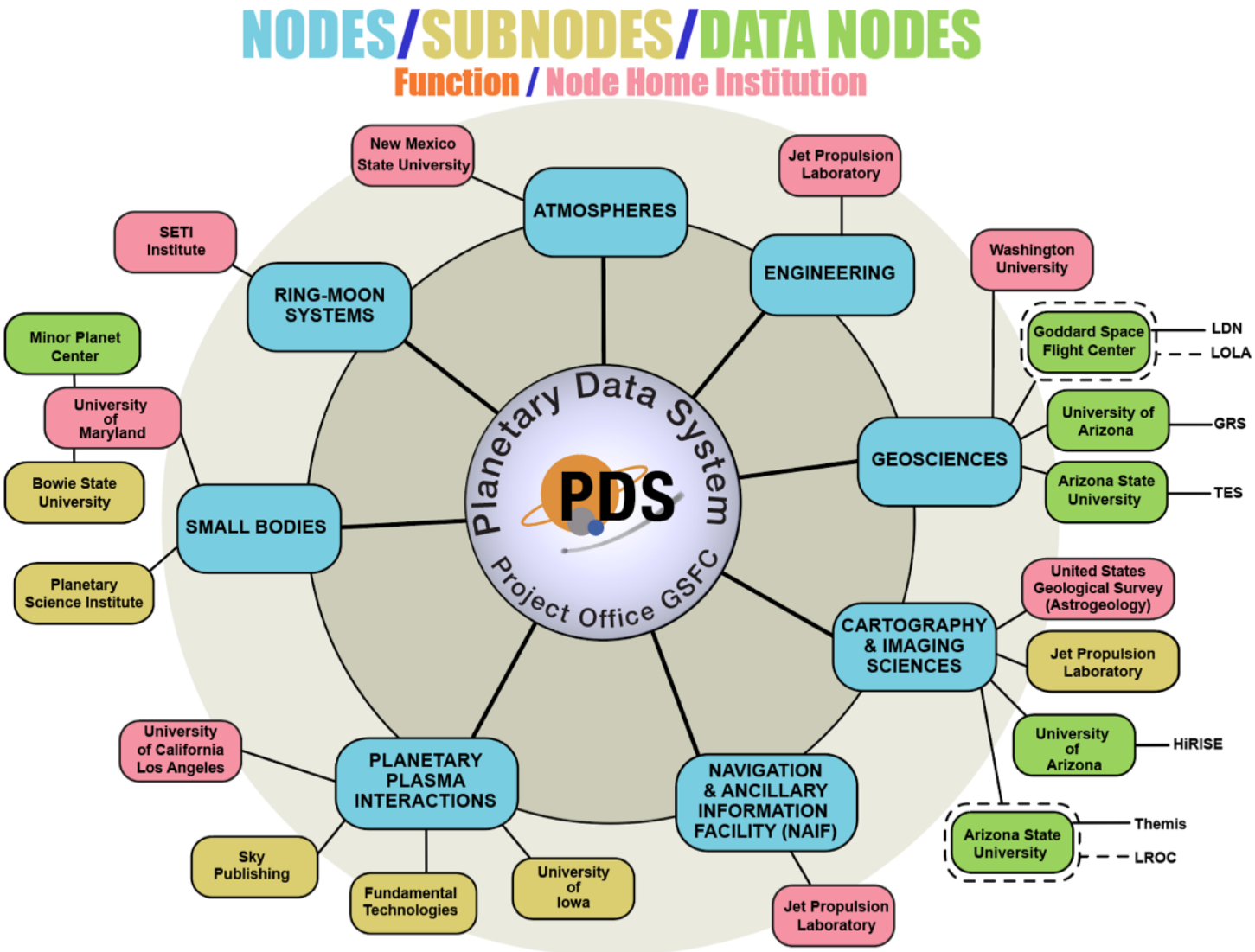
- PDS works with every NASA planetary mission with the goal of obtaining a complete archive of data from that mission.

- PDS4 is now an international standard accepted by almost all space faring nations through the offices of the COSPAR Sponsored International Planetary Data Alliance (IPDA).

Overview - PDS Structure Detailed



Overview - PDS Structure Detailed



Overview - Who are we?

Atmospheres Node:
New Mexico State University,
Las Cruces, New Mexico
(Nancy Chanover, Manager)

Geosciences Node:
Washington University,
St. Louis, Missouri
(Ray Arvidson, Manager)

Cartography & Imaging Sciences Node:
US Geological Survey,
Flagstaff, Arizona
(Lisa Gaddis, Manager)

Planetary Plasma Interactions Node:
University of California,
Los Angeles, California
(Raymond Walker, Manager)

Ring-Moon Systems Node:
SETI Institute,
Mountain View, California
(Mark Showalter, Manager)

Small Bodies Node:
University of Maryland,
College Park, Maryland
(J. (Gerbs) Bauer, Manager)

Discipline Nodes

Support Nodes

Engineering Node:
Jet Propulsion Lab, Pasadena, California
(Dan Crichton , Mgr.)

Navigation and Ancillary Information Facility:
Jet Propulsion Lab, Pasadena, California
(Chuck Acton, Mgr.)

William Knopf
Program Executive, NASA Headquarters

Micheal New
Program Scientist, NASA Headquarters

Ralph McNutt
PDS Chief Scientist, JHU Applied Physics Lab

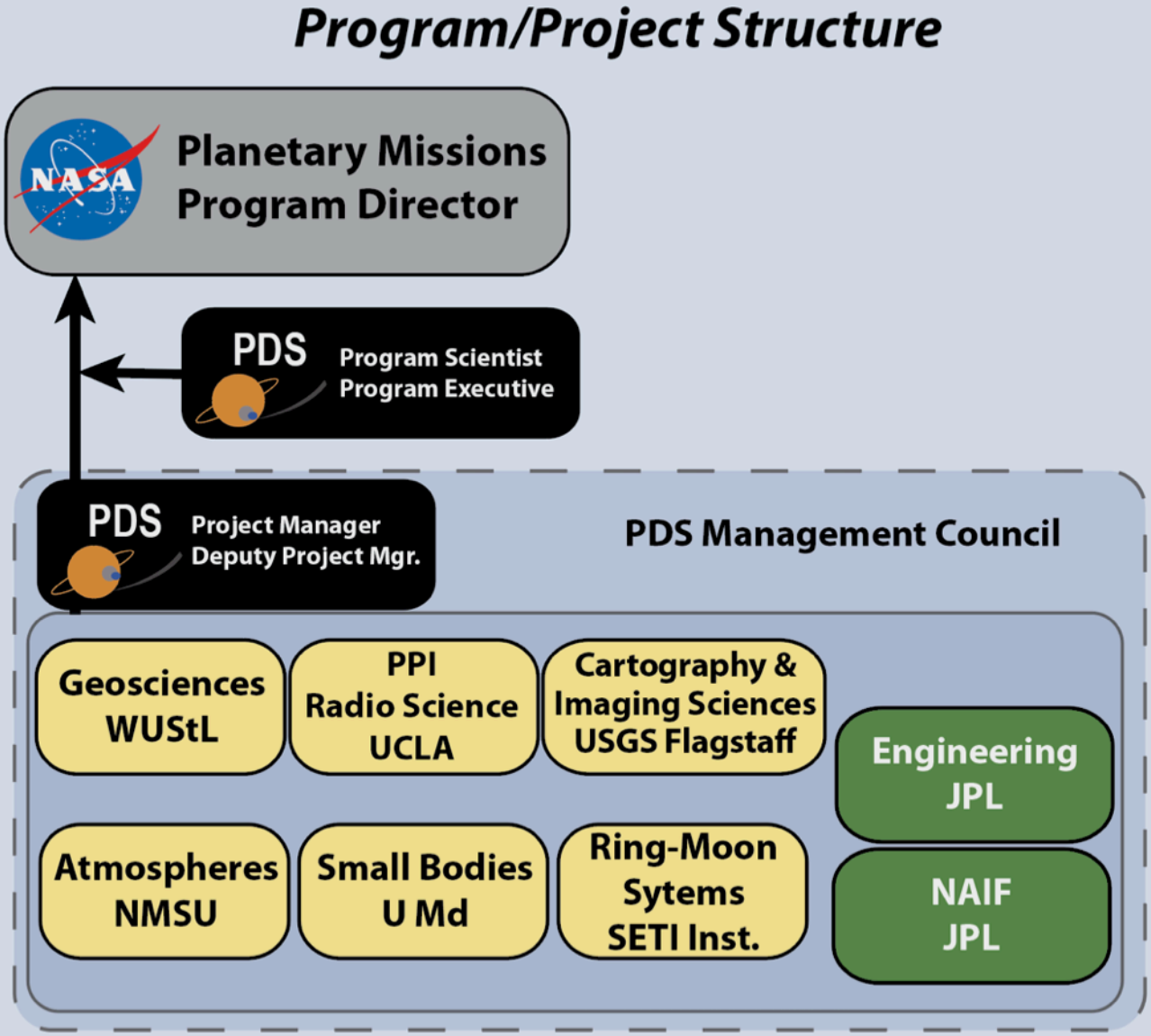
Tom Morgan
Project Manager, NASA GSFC

Management

Overview - Who are we?

Images of some of the more than 70 scientists, engineers, archivists, and IT professionals who make the PDS work.

Overview - How we are managed



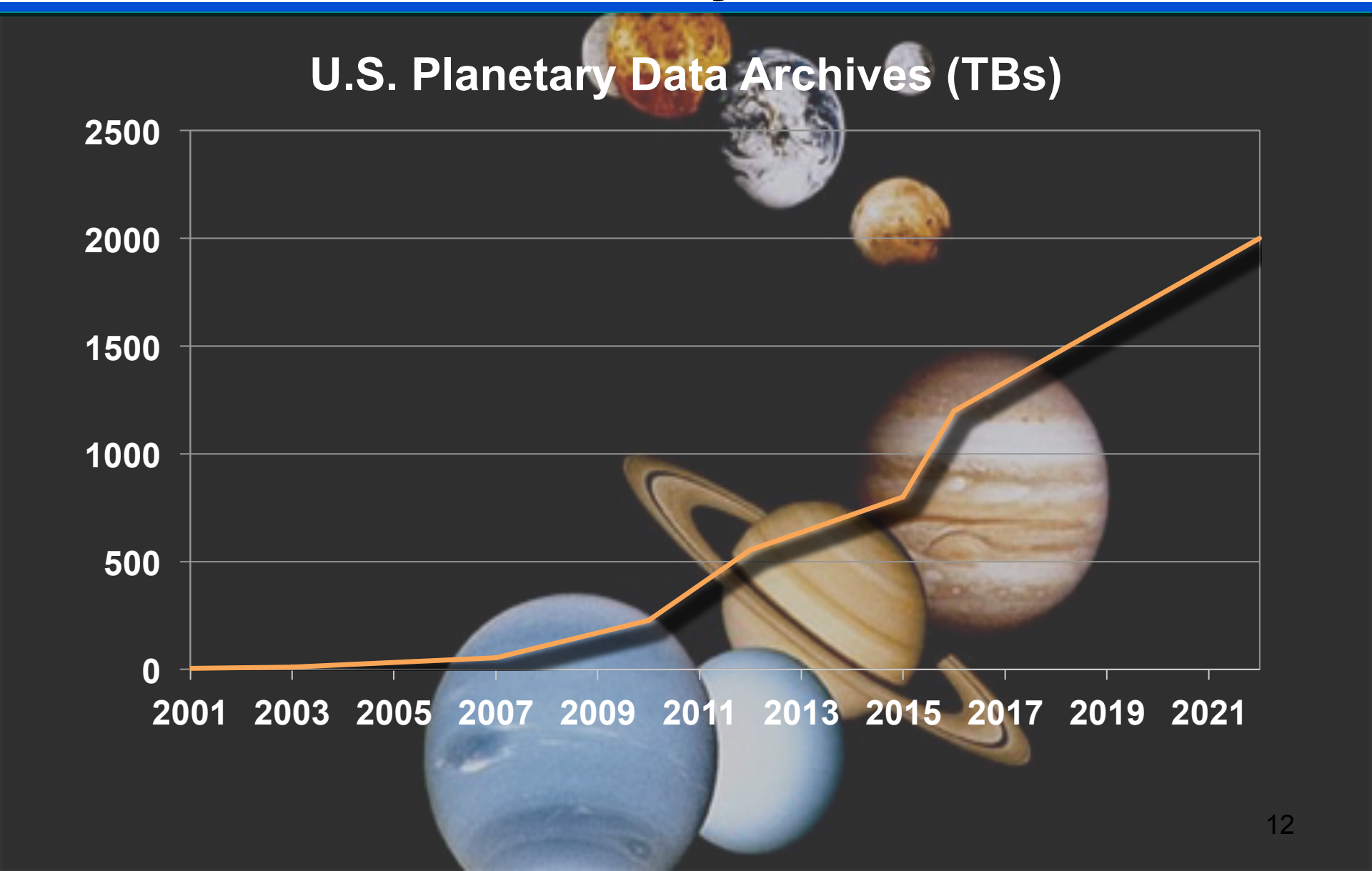
PDS is a federated system with a **Management Council** made up of:

- Node leads
- Project Manager
- Deputy Project Manager

This group serves as the technical authority for the PDS

Per NNH15ZDA006C (PDS CAN) and PDS Project Plan

Growth of Planetary Data Archived from U.S. Solar System Research



Overview - Scale and Diversity of the PDS

Web statistics:

Jan – Dec, 2017	All Nodes
U.S. Files	78,591,894
Intl. Files	21,924,844
U.S. Volume	341,034,526
Intl. Volume	45,988,420
U.S. Visitors	58,230
Intl. Visitors	19,304

- **PDS contains ~1.4PB of data from >70 missions.**
- **Represents >2000 data sets from >600 unique instruments**
- **Thirteen missions are currently delivering data to the PDS.**

Active US Missions - Stoplight Chart

Active US Missions	Budget				Schedule				Technical				Mission				Node
Jan-18	*Oct*	*Nov*	*Dec*	*Jan*	*Oct*	*Nov*	*Dec*	*Jan*	*Oct*	*Nov*	*Dec*	*Jan*	*Oct*	*Nov*	*Dec*	*Jan*	
2001 Mars Odyssey	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	GEO
2003 Mars Exploration Rovers	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	GEO
2005 Mars Reconnaissance Orbiter	G	G	G	G	Y	Y	Y	Y	G	G	G	G	Y	Y	Y	Y	GEO
Lunar Reconnaissance Orbiter	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	GEO
Mars Science Laboratory	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	GEO
InSight (PDS4)	G	G	G	G	G	Y	Y	G	G	G	G	G	G	Y	Y	G	GEO
Mars 2020 (PDS4)	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	GEO
Cassini	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	ATM
New Horizons	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	SBN
Dawn	G	G	G	G	Y	Y	Y	Y	G	G	G	G	Y	Y	Y	Y	SBN
MAVEN (PDS4)	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	PPI
Europa Clipper (PDS4)	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	CIS
OSIRIS-Rex (PDS4)	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	SBN
Juno	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	ATM
Lucy (PDS4)	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	SBN
Psyche (PDS4)	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	SBN
DART	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	SBN

Active US Missions – Mission Data Status Update Details

DEFINITIONS
GREEN: There are no current problems and none if foreseen in meeting commitments.
YELLOW: Problems exist; solutions have been identified, but one or more commitments are in jeopardy. At presents, the problems are being managed within the DN, but close monitoring is required.
RED: One or more major problems exist with no identified solution; the associated commitments will not be met.

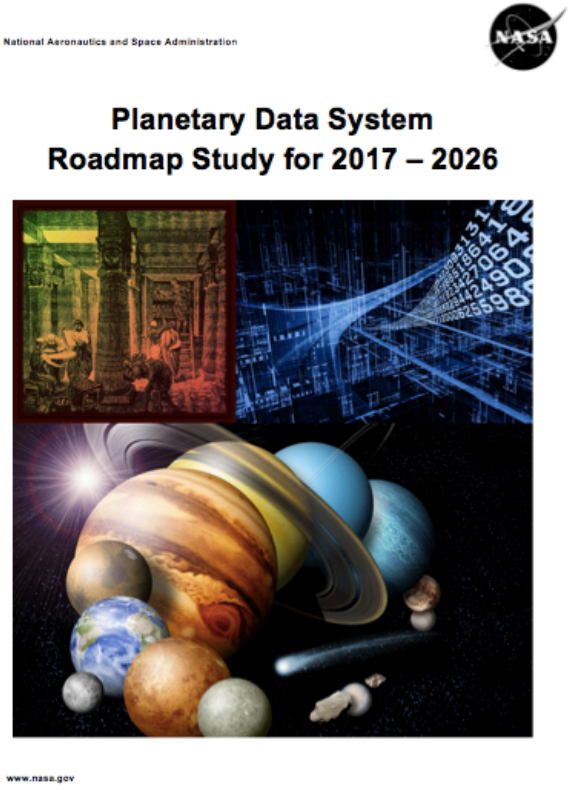
Dawn: VIR (Ceres) (YELLOW)
Dawn VIR is yellow because it is still behind the planned schedule, although making very good and steady progress. We are now expecting VIR archiving to be completed sometime in early 2018.
Details: The data now in liens resolution include L1a and L1b for PM (Prime Mission) LAMO and first half of EM1 (First Extended Mission). The data being prepared for upcoming review include L1a and L1b through the end of EM1, Ceres global maps, and Ceres L1c (artifact-corrected data, an extra product).

MRO: SHARAD (YELLOW)
The SHARAD team has resumed deliveries from the ASI (Italian Space Agency) SHARAD Ground Operations center, to recover the backlog of data from deliveries that were missed due to a hiatus in ground operations between June 2012 and March 2013. The Geoscience node has received, validated and posted updates to some ancillary SHARAD data files from ASI. The team expects to be caught up with the backlog of EDR deliveries in time for Release 44, 3/1/18.

What was the PDS Roadmap Study?

The PDS Roadmap Study was chartered in 2015 by the Planetary Science Mission Division, following the release of an RFI for information from the Planetary Science Community.

- Dr. Ralph McNutt was appointed the chair of the Roadmap team.
- Report was released in July 2017.
- Results of the Roadmap were cast as “Findings.” There were 19.
- The report can be found at https://pds.nasa.gov/roadmap/index.shtml.



Purpose: The new Roadmap was to chart the future course of the PDS for the wide variety of PDS stakeholders (including data archivists, data providers from large NASA missions to small research investigations, data users, NASA managers, PDS staff, and the general public) for the next decade and beyond.

Establishing the Roadmap Activity

The Roadmap Study Team (RST): 16 individuals with differing backgrounds and interactions with the PDS

Discussions began with the items in the RFI, namely, consideration of:

- ✓ Tools, resources, workflows, tutorials, and interfaces
- ✓ Making the archiving process seamless, less costly, and more efficient
- ✓ Role of PDS relative to other archiving alternatives (i.e. journals), in providing public access to NASA-generated data
- ✓ Integration of PDS data products and services with those of other facilities (e.g., U.S. Geological Survey's cartography program, Minor Planet Center)
- ✓ Role of PDS in encouraging development of higher-order data products
- ✓ Appropriate improvements to current search capabilities of the PDS

Results of Roadmap Activity

The final Roadmap report was the RST's understanding of the PDS as it now operates, as well as findings of fact, and suggestions for goals for the PDS and NASA to consider in furtherance of PDS activities during the upcoming decade.

The results were captured in **19 findings**:

These include:

- Findings of fact that require no action (10)
- Findings that require action primarily at the program level (HQ) (2)
- Findings that require action, but that the PDS can directly address (5)
- Findings that require both program and project level action (2)

The sorting below is that of the project office – not the Roadmap team - and is based in part on what we think we can do in the near future.

Full List of PDS Roadmap Findings

Red = Program Level

Blue = PDS Project Level

Black = Program/Project Level

Green = no action required (colors are mine)

The color coding is not that of the RST. It is used here simply for purposes of organizing this presentation.

Subjects of the Findings:

1. PDS Stakeholders.
2. Managing Expectations of PDS Usability.
3. Data Discoverability.
4. Integration with Other Archives.
5. Citation of Data Sets.
6. Modernizing Metadata.
7. Access to Data
8. Documentation and Training..
9. PDS File Formats and Translation Software.
10. Archiving Software.
11. Information Technology.
12. Potential Impact of ROSES Archiving Requirements.
13. Higher-Order Data Products.
14. Astromaterials Data I.
15. Astromaterials Data II.
16. Node Organization.
17. Transparency.
18. PDS Governance.
19. Timing of the Next PDS Roadmap Study.

Addressing the Roadmap findings
Project Office *Approach and Implementation*

- **We can't address address every finding at once.**

- Identify (with HQ and MC) high priority Roadmap Findings that are within PDS' purview with additional funding (7 of the 18 findings).

- Start with what we have already: *i.e.* what tools do we already have or that are already in progress.

- Work through the MC and our nodes through an iterative process to identify additional tasks needed that require additional funding. Targeted tasks include those that can be initiated in the next 2 years:
  - Each node asked for input and to identify needed tasks with effort estimates.
  - Management Council deliberations arrived at cross-PDS consensus of top priority tasks.

Addressing the Roadmap findings

Identifying the "High Priority" findings

3. Data Discoverability. There is a need for PDS to both expand and focus its search services, with a view to making it easier for users to find and execute the search appropriate to their query.

5. Citation of Data Sets. PDS is actively involved in addressing the data citation issue, and is well-positioned to provide the essential links in the chain needed to enable clear, direct referencing of PDS products; but it cannot itself change the habits and attitudes of authors, referees, and journal editors when it comes to including data set references in publications.

6. Modernizing Metadata. The accessibility and discoverability through the PDS4 metadata registry is a cornerstone to the future of community interaction with the PDS as a coherent storehouse of data. Legacy data archived in PDS3 format (the vast majority of PDS holdings) often lack metadata sufficient to enable discovery and accessibility commensurate with PDS4.

Red = Program Level; Blue = Project Level; Black = Program/Project Level; Green = no action required (PO assignment of colors)

8. Documentation and Training. The PDS4 information model is well-documented at a highly technical level. However, there is a critical need for broader documentation and training for all levels of users..

9. **PDS File Formats and Translation Software. There is a need for more translation programs that transform data from the PDS4 archive file formats to more usable analysis-ready formats.**

17. Transparency. The use and application of PDS4 standards and development of third-party support for PDS4 metadata and formats is hindered by a lack of transparency in the PDS development process.

18. **PDS Governance. NASA management has not settled the question of how PDS fits into current NASA governance structures. PDS has a minimal Project Office, which lacks resources for providing detailed cross-discipline reports, studies, and guidance as there are within other NASA SMD data activities, which would put a more unified public face on the PDS and support other activities provided for in the current NASA governance model.**

Addressing the Roadmap Findings

The Roadmap was released on June 23, 2017

Both the PDS MC and the Planetary Science Division (PSD) were given interim reports on the progress of the Roadmap.

The Roadmap Chairman made a formal presentation to PSD leadership, September 5, 2017:

The Program Manager (David Schurr) directed the PDS to develop a plan to allocate a 10% augmentation of current PDS efforts (~7 WYE equivalent per year) over the next 2 years to address *high priority roadmap issues.*

HQ identified several key areas to be *immediately addressed* including:

1) Completing a project plan (Finding #18)
2) Increased staffing of the project office (Finding #18)
3) Training for archiving in PDS4 focused on small providers (Finding #8)
4) Accelerating PDS3-PDS4 translation (Finding #6)

Addressing High Priority Roadmap Findings

Immediate Actions

Actions Already initiated:

- ✓ Finding #18: The Draft PDS project plan was submitted to PDS Nov. 17th 2017.
- ✓ Finding #18: Code 690.1 (GSFC) is now advertising a new CS position for the project office; the position opening will be posted on Feb. 28.
- ✓ Finding #8: We brought in a part-time contractor for 3 months to serve as a PDS Training and Communication Coordinator to identify the needs of the PDS and community.
  - Depending on the final 2018 budget, we expect to seek a full-time PDS Training and Communication Coordinator.
- ✓ Finding #6: PDS3-PDS4: Additional funding has been allocated for $1.2M for key PDS3-PDS4 migration efforts. We are now in the process of distributing these funds to the nodes.

Planning for the next two years:

- ✓ The PDS collectively, over 2 F2F meetings and through several PDS wide working groups (tools working group, web design working group, training working group), developed a draft proposal to address the Roadmap findings and presented it to PSD senior management January 8, 2018.
- ✓ Invited back post CR to discuss additional funding.

Finding #8 Immediate Actions

PDS4 Training: 3rd Planetary Data Workshop 2017

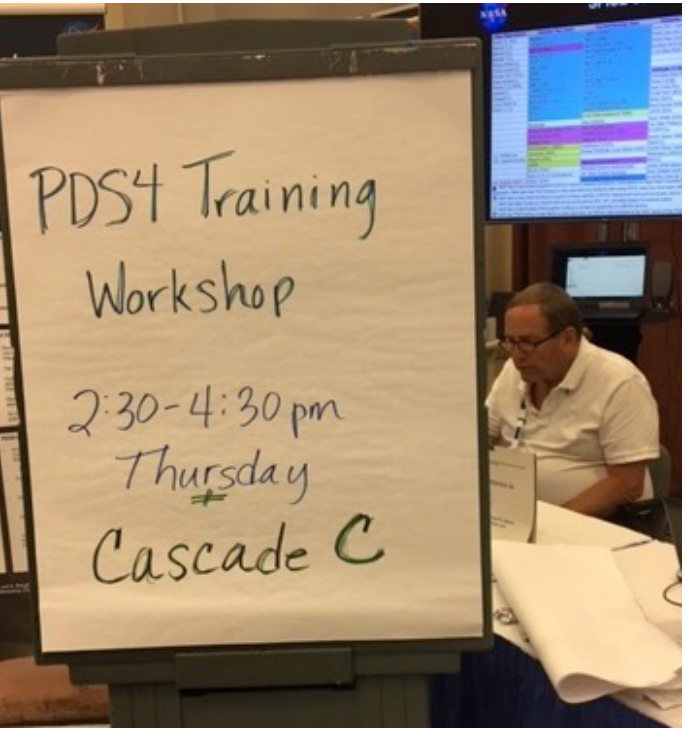
1. **Workshops for new users with hands on training**
   - Large Classroom style
2. **Cross-node support**

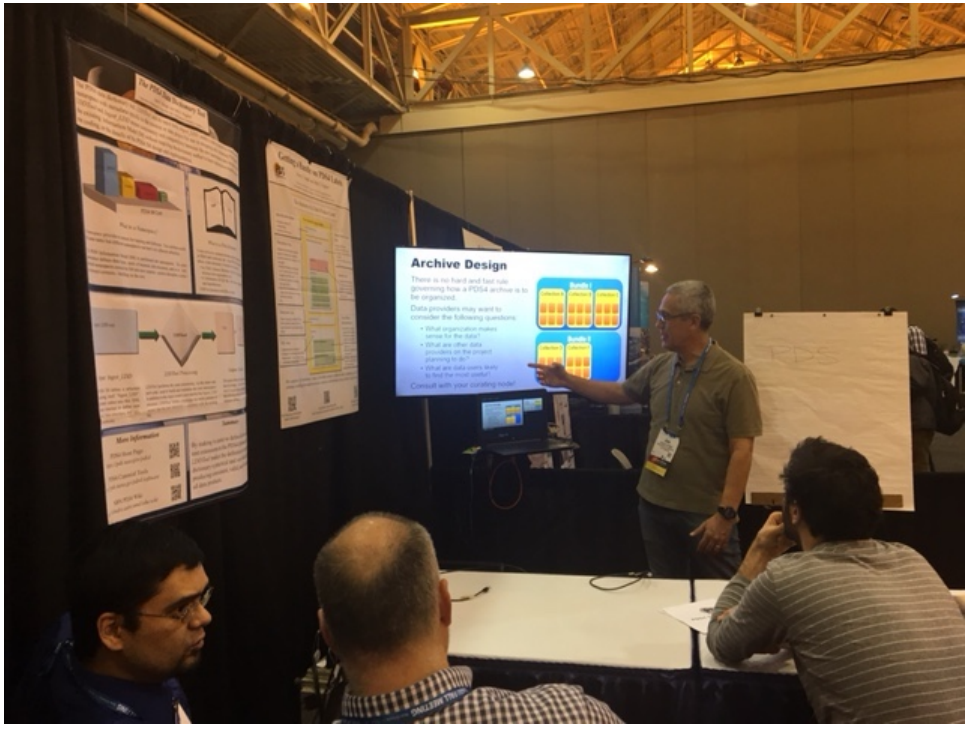
Addressing the Roadmap Findings

Finding #8 Immediate Actions

PDS4 Training: DPS Meeting 2017

1. Two training workshops: workshops for new users
   - Small classroom style
2. One on one training
3. ATM lead this effort with additional cross-node support





Addressing the Roadmap Findings

Finding #8 Immediate Actions

PDS4 Training: AGU 2017

1. One-on-one training at a booth
2. PPI node leading with additional cross-node support

Addressing the Roadmap Findings

Finding #8 Planning for the next 2 years

PDS4 Training: The main path forward:

Create training modules that can be used in multiple training situations

1. **Develop an inward facing website where modules can be accessed, downloaded, edited and uploaded by all of the node trainers and developers**
2. **Collaboration between nodes for training development and events**

Long Term Goals

1. **Create an outward facing website where the individual user is able to access modules and videos to self teach**
2. **Determine and implement most effective methods for training**
   - Booth vs. classroom
   - Scientific conferences vs. PDS4 retreat

Addressing the Roadmap Findings

Finding #8 Planning for the next 2 years

PDS4 Training and Communication Coordinator: Position overview **(1.0 FTE for two years)**

- **Attend meetings/conferences with training events**
  - Ensures uniformity and continuity between training events
  - Gather feedback from training participants: track training wants, needs, and effectiveness
- **Regular communication with training team: coordination with nodes**
- **Develop/coordinate training materials:**
  - PDS4 training modules and videos: subjects have been selected.
  - Training websites

Addressing the Roadmap Findings

Finding #8 Planning for the next 2 years

- **Lunar and Planetary Science Conference, 2018**
  - Larger booth area for individual/small group hands on training
  - Lead by GEO node with additional cross-node support
- **Europa Clipper Mission team**
  - June 2018
  - Lead by CIS node

PDS4: The Next Generation

PDS-wide project to upgrade from PDS version 3 (PDS3) to address many of the big data challenges of a large-scale, distributed, international system

PDS4 is model-driven, service-oriented, and XML-based:

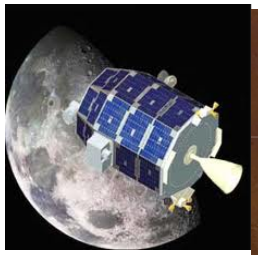
- Provides a comprehensive global information model and hence a unitary archive metadata architecture.
- A modern technical foundation for planetary science data.

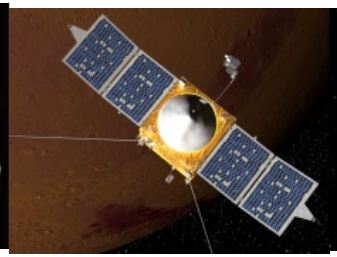
PDS4: The Next Generation

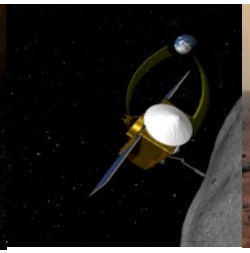
- **An explicit information architecture**
  - **All products are tied to a common model** for validation and discovery
  - **Use of XML**, a well-supported international standard, for labeling, validation, and searching
  - A hierarchy of dictionaries built to the ISO 11179 standard, designed to increase flexibility, enable complex searches, and make it easier to share data internationally
- **Distributed services both within PDS and at international partners**
  - Consistent protocols for access to the data and services
  - Deployment of an open source registry infrastructure to track and manage every product
  - A distributed search infrastructure

Addressing the Roadmap Findings

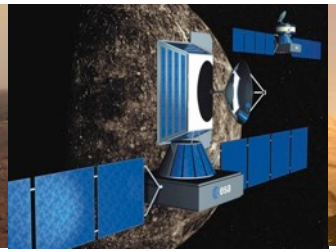
Finding #6 Already in Motion


LADEE (NASA)


MAVEN (NASA)


Osiris-REx (NASA)


ExoMars (ESA/Russia)

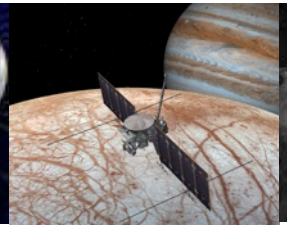

BepiColombo (ESA/JAXA)


Mars 2020 (NASA)

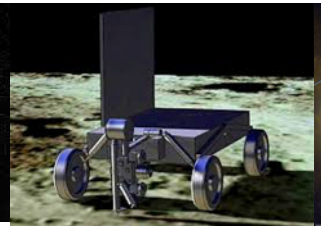

Psyche (NASA)

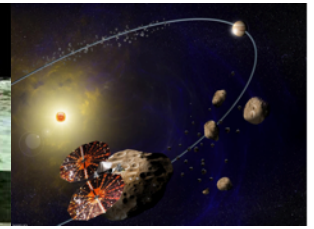

InSight (NASA)


JUICE (ESA)


Europa (NASA)


Hyabussa-2 (JAXA)


Chandrayaan-2 (ISRO)


Lucy NASA

Endorsed by the **International Planetary Data Alliance** in July 2012
https://planetarydata.org/documents/steering-committee/ipda-endorsements-recommendations-and-actions

Addressing the Roadmap Findings

Finding #6 Immediate Actions

PDS3 to PDS4 Translation of legacy data: Highest Priority Missions

Mission	Delta (2018)	Breakdown by node
Grand Total	**$1135.2k**	
Cassini	$652.7k	• Planetary Plasma Interactions • Atmospheres • Cartography & Imaging Sciences • Ring-Moon Systems
MESSENGER	$307.5k	• Planetary Plasma Interactions • Cartography & Imaging Sciences • Geosciences
Dawn	$100.0k	• Small Bodies: Planetary Science Institute
Rosetta (US-team Instruments)	$75k	• Small Bodies: University of Maryland

Addressing the Roadmap Findings

Finding #6 Planning for the next 2+ years

Remaining High Priority Legacy Data Sets Identified for translation to PDS4 through 2021:

Additional estimated funds needed for Effort: **~$1,327.0k**

(data sets & est. funding from spring 2017, to be updated for spring 2018)

- New Horizons (Pluto and Jupiter only)
- Voyager 1 and 2 (Jupiter, Saturn, Uranus, Neptune)
- Clementine LIDAR, Lunar Topo, LWIR, Gravity
- GRAIL Cal/Resampled Science, Derived Gravity Science Data Products
- Magellan
- Mars Global Surveyor MOLA, TES
- Phoenix MECA, TGA
- Viking Lander 1 & 2 LCS, Labelled Release, Seismology.
- Huygens ACP, DISR, DWE, GCMS, HASI, SSP
- MSL REMS
- Juno
- Galileo

Addressing the Roadmap Findings

Finding #6 Planning for the next 2+ years

Approach for Active Missions: Considerations for scoping the translation effort for a given instrument team

- What data should be translated?
  - All PDS3 data sets for the instrument?
  - Ancillary data: browse, calibration, extras, etc.
- For each of the above, is it A, B, or C and what work is involved?
  - A = PDS-compliant data, PDS3 keywords easily mapped into a PDS4 label.
  - B = PDS-compliant data, some dictionary work needed to make a PDS4 label.
  - C = Data are not PDS-compliant.
- What software will need to be acquired, revised, or created?
  - Tools to design and mass-produce PDS4 labels
  - Revisions to existing software that reads PDS3
- What documentation will need to be revised or created?
- Who is best positioned to do the work?
  - Instrument team
  - PDS node
  - Third party such as MIPL/ OPGS

Geosciences Node efforts in helping active missions make their translation plans

ODY	MER	MRO	MSL	LRO	Planning steps
✓	✓	✓	✓	✓	Geo opens conversation with mission and offers to help make plan
✓	✓	✓	✓	✓	Geo creates preliminary assessment of level of effort for each data set, with input from the other nodes archiving for the mission
✓	2/20/18	✓	✓	✓	Geo and mission hold initial DAWG meeting so all teams and nodes are informed
IP		IP	IP	IP	Nodes and individual teams meet to scope the work for each data set
IP		IP	IP	IP	Nodes and teams break down conversion work into steps and assign each step to team, PDS, or others (e.g. OPGS/MIPL)
IP		IP	IP	IP	Responsible parties estimate the costs of each step
					Mission compiles the results in a report to deliver with extension proposal in April 2018

Addressing the Roadmap

Plans for the next Two+ Years

NEED	Rationale	Yearly Commitment
Website Redesign. *Goals:* -Improve Data Discoverability: expand and deepen data search services -Make PDS organization/tools/ information transparent to users -Consistency across node webpages	**Finding #3:** The website design is 10 years old.	0.5 WYE for two years. • Could be implemented by RFP or by assignment to a NASA center (PM thinks this is going to cost us more).
Tool Development and cross-node coordination. *Goals:* - Increase the number and reliability of tools and extend functionality for cross-node use -Coordinate PDS-wide efforts and reduce redundancy. -Stand alone software with good documentation after 2 yrs	**Finding #9:** More programs that transform data from the PDS4 archive file formats to more usable analysis-ready formats are needed. Nodes developed ingest tools for their local use. Improvements and PDS wide beta testing are needed to complete tools useable across the PDS and community.	3.0 WYE for each of two years. • 1.0 WYE for cross-node coordination/ monitoring and for node beta testing • 2.0 WYE for high priority tool development/ extension of functionality)
PDS3-PDS4 Cross node coordination	**Finding #6**: Systemized and coordinated conversion of data	0.7 WYE for each of 2 years only.

Addressing the Roadmap

Plans for the next Two+ Years

NEED	Rationale	Yearly Commitment
Training coordinator. *Goals:* • Coordinate cross-node training efforts, materials, and documentation (i.e. PDS4 archiving, DMP development) • Develop training module materials and design uniform PDS educational/outreach materials	**Findings #8 and #12:** • New ROSES archiving requirements. • There is a critical need for broader documentation and training for all levels of users.	1.0 WYE, for each of 2 years. • We hope to have enough progress to make training largely web-based by the end of 2-years.
External measures of customer satisfaction. *Goal:* Provide an independent printed annual assessment of the PDS by node.	**Findings #18:** • Minimal Project Office • Resources are needed to provide detailed cross-discipline reports, studies, and guidance.	0.4 WYE, ongoing, plus survey cost.
Cyber Security	We have performed our first POE&M, but there are costs associated with implementing CATS for the nodes.	0.7 WYE plus procurements, ongoing • we have about a $100,000 backlog.

Conclusion and Future Plans

Despite the CRs, in the 8 months since the Roadmap was released, we have been able to begin addressing key Roadmap findings.

Into the future we hope, with the support of the PSD, to:

- continue the work on PDS3-PDS4 translation and training
- expand tool development and cross-node tool coordination
- improve web design
- develop actionable external measures of customer satisfaction

There are key findings that need Program level attention (Findings #1 and #2 – see detailed description in backup slides) .

The final 2 slides are:

1) an add for an important technical opportunity sponsored by the PDS, that some of you may wish to attend
2) our current level 1 requirements

On behalf of the PDS, thank you for hearing us!

1st Planetary Science Informatics & Data Analytics Conference

https://psida.rsl.wustl.edu

Data Management, Cloud Computing, Data Modeling, Information Systems, Information Processing, Machine Learning, Visualization, and more…

Current Mission and Level 1 Requirements

PDS Mission Statement

The mission of the Planetary Data System (PDS) is to facilitate achievement of NASA's planetary science goals by efficiently collecting, archiving, and making accessible digital data and documentation produced by or relevant to NASA's planetary missions, research programs, and data analysis programs.

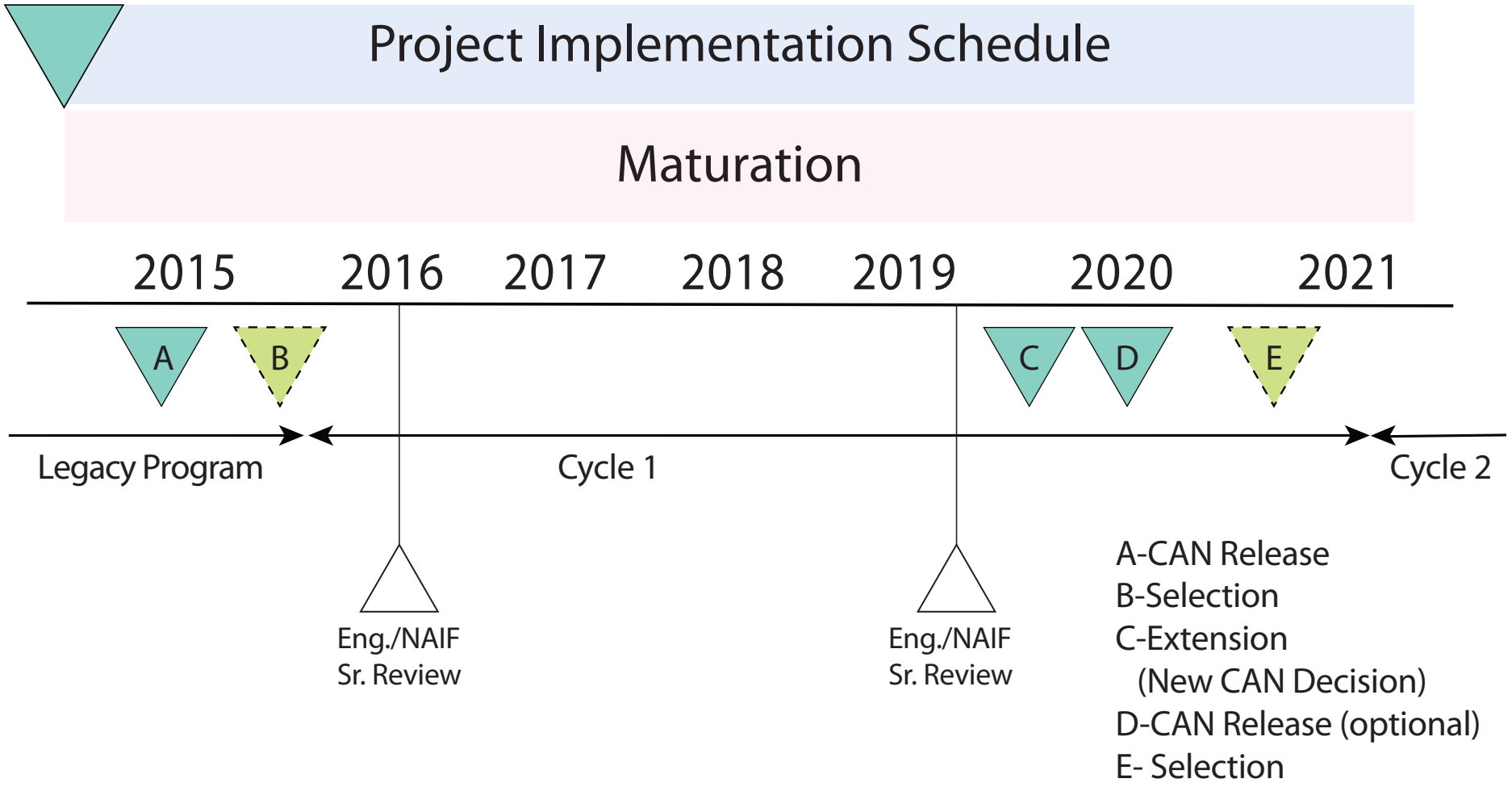
Level 1 Requirements

✧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.
✧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.

Level 1 requirements are levied by Headquarters.

Backup Slides

PDS Project Schedule



Archive Size by node

PDS contains 1.4PB of data from more than 70 missions.

Thirteen missions are currently delivering data to the PDS.

Web Statistics are shown below

2017 Jan.-Dec.	PDS	ATMOS	EN	GEO	IMG	NAIF	PPI	Rings	SBN-PSI	SBN-UMD
U.S. Files	39,295,947	1,223,637	3,675,090	17,348,600	10,934,755	4,041,919	376,039	1,338,444	39,562	317,901
U.S. Volume	170,517,263	2,940,816	489,045	33,025,942	115,518,661	11,276,337	2,118,818	893,534	1,378,974	2,875,136
U.S. Visitors	29,115	644	7,295	3,769	4,533	8,854	1,122	1,004	773	1,121
Intl Files	10,962,422	13,365	37,591	1,056,213	8,994,803	485,152	301,908	27,979	41,921	3,490
Intl Volume	22,994,210	3,133	29,574	6,597,218	7,337,577	4,076,037	3,799,643	3,007	1,119,257	28,764
Intl Visitors	9,652	155	2,110	1,798	2,228	1,490	282	760	324	505

Establishment of the Roadmap Activity Details

The Roadmap activity was established to look ahead to the time period from 2017 to 2026. The Roadmap activity began on 5 November 2015 with the release of a Request for Information (RFI) asking for community input by 5 January 2016.

NASA chartered this current PDS Roadmap Study Team to "develop a practical, community-developed pathway to implement the new long-term vision for the PDS, which continues to accomplish NASA's broad objective for the PDS; namely, preserving and making available all data products from planetary exploration research and missions."

The New Roadmap was to chart the future course of the PDS for the wide variety of PDS stakeholders (including data archivists, data providers from large NASA missions to small research investigations, data users, NASA managers, PDS staff, and the general public) for the next decade and beyond.

Full List of PDS Roadmap Findings with details

Red—Program Level; Blue PDS Project Level; Black--Program/Project Level; Green—no action required (colors are mine)

1. **PDS Stakeholders. While all PDS stakeholders are recognized as valuable, the prioritization of stakeholder interests and the impact those interests should have on PDS policy, design, and resource allocation are unclear.**

2. **Managing Expectations of PDS Usability. There is a mismatch between the services and functions PDS is equipped to provide and the very high expectations of its users and NASA management.**

3. Data Discoverability. There is a need for PDS to both expand and focus its search services, with a view to making it easier for users to find and execute the search appropriate to their query.

4. Integration with Other Archives. The PDS serves as the model for other national space-mission data archives in ensuring future universal accessibility and searchability. The PDS is uniquely poised to lead efforts to make national and global archives interoperable.

5. Citation of Data Sets. PDS is actively involved in addressing the data citation issue, and is well-positioned to provide the essential links in the chain needed to enable clear, direct referencing of PDS products; but it cannot itself change the habits and attitudes of authors, referees, and journal editors when it comes to including data set references in publications.

6. Modernizing Metadata. The accessibility and discoverability through the PDS4 metadata registry is a cornerstone to the future of community interaction with the PDS as a coherent storehouse of data. Legacy data archived in PDS3 format (the vast majority of PDS holdings) often lack metadata sufficient to enable discovery and accessibility commensurate with PDS4.

7. Access to Data. The PDS does an excellent job of providing access to its data holdings and is on track to increase such access. The latter is enabled by the PDS4 uniform metadata standard.

8. Documentation and Training. The PDS4 information model is well-documented at a highly technical level. However, there is a critical need for broader documentation and training for all levels of users..

9. **PDS File Formats and Translation Software. There is a need for more translation programs that transform data from the PDS4 archive file formats to more usable analysis-ready formats.**

10. Archiving Software. The PDS is not an appropriate archive or repository for software.

Full List of PDS Roadmap Findings with details continued

11. Information Technology. The PDS has been and continues to be proactive in investigating information technology and adopting best practices.

12. **Potential Impact of ROSES Archiving Requirements. It is a matter of concern as to whether the PDS nodes will have the resources to serve the data archiving requirements of individual ROSES investigations.**

13. **Higher-Order Data Products. Higher-order products produced by mission teams beyond what is in their original data management plans are extremely valuable additions to the archive but are not always included due to lack of resources needed by missions to complete the archiving process.**

14. **Astromaterials Data I. A large amount of data from laboratory analyses of samples obtained by NASA missions is not archived and is in danger of loss. Astromaterials data today is primarily stored on short-lived media, in private holdings, and with PI dependent documentation.**

15. **Astromaterials Data II. A large amount of data from laboratory analyses of meteorites and cosmic dust is not archived and is in danger of loss.**

16. Node Organization. PDS funding levels, combined with the lack of opportunity to propose new nodes separate from the re-compete activity for existing nodes, has had the effect of strongly discouraging the establishment of new nodes or otherwise allowing the PDS organization to grow to keep pace with development and expansion of Planetary Science disciplines and technology.

17. Transparency. The use and application of PDS4 standards and development of third-party support for PDS4 metadata and formats is hindered by a lack of transparency in the PDS development process.

18. **PDS Governance. NASA management has not settled the question of how PDS fits into current NASA governance structures. PDS has a minimal Project Office, which lacks resources for providing detailed cross-discipline reports, studies, and guidance as there are within other NASA SMD data activities, which would put a more unified public face on the PDS and support other activities provided for in the current NASA governance model.**

19. **Timing of the Next PDS Roadmap Study: This Roadmap Study was initiated in the year immediately following a competition of the PDS Nodes, and will be completed at least 3 years (and perhaps longer) before the next recompetition, which limits the impact of a Roadmap Study activity on shaping the work of the PDS.**

Red—Program Level; Blue PDS Project Level; Program/ Project Level; Green—no action requested

Addressing the Roadmap Findings

Finding #6

Finding IX: There is a need for more translation programs that transform data from the PDS4 archive file formats to more usable analysis-ready formats.

A suggested action to address this finding is for the PDS to support format translation from PDS4 format to a few selected file format standards, and also to provide encouragement and support to non-PDS software developers to include PDS4 input/output capabilities in their tools. PDS has made a good beginning to this task but more work remains.

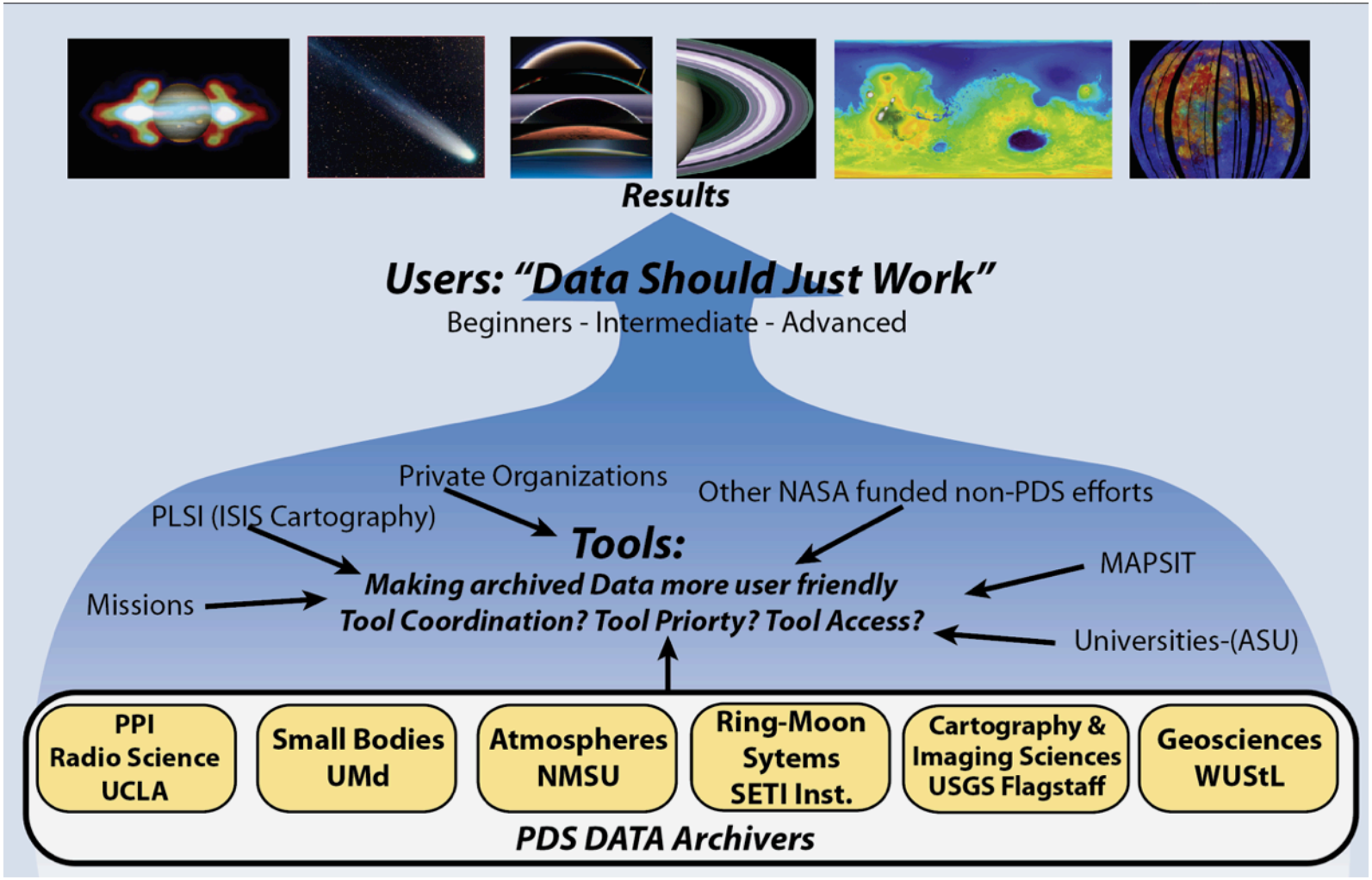
Immediate Response

The PDS science nodes are working with their user groups to identify analysis ready formats.

This is part of a larger cluster if issues. See the next slide.

Addressing the Roadmap Findings

Finding #9: The big Tools Picture



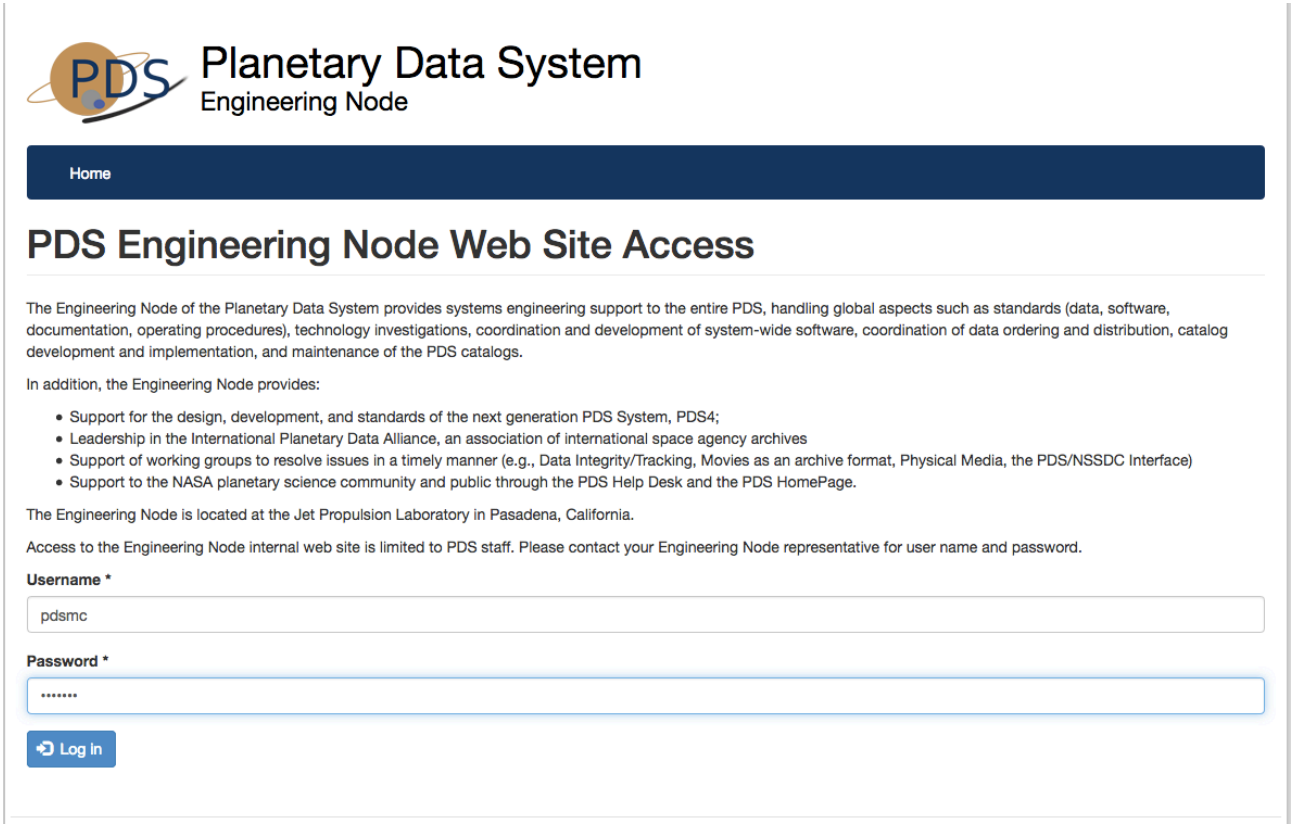
We are not alone.

Addressing the Roadmap Findings

Finding #16

The problem: much of the information a software developer would require was behind a protected site.

Solution: Redesign the web site.



Addressing the Roadmap Findings

Finding #16

Solution continued:
Standards information is now online at https://pds.nasa.gov/pds4/doc which includes standards documents and schemas along with information about changes for each release.

PDS will post information about the change and release process. This will include contact information along with information on how to submit change requests.

PDS will include a release timeline in reverse chronological order with the top date identifying the next release target and change information for each release is highlighted.

Longer term, PDS will provide information on use of standards for developers including how to use JSON and RDF extractions of the information model to develop software.

This will be part of a larger effort to redesign the website.