PDS4 Overview and Status

https://pds.nasa.gov/

Dan Crichton
Manager, PDS Engineering Node
Jet Propulsion Laboratory

June 13, 2017

1

Planetary Data System

- <u>Purpose:</u> To collect, archive and make accessible digital data and documentation produced from NASA's exploration of the solar system from the 1960s to the present.

- <u>Infrastructure:</u> A highly distributed software infrastructure with planetary science data repositories implemented at major government labs and academic institutions
  - System driven by a well defined planetary science information model
  - Over 1 PB of data
  - Movement towards international interoperability through IPDA
  - Distributed federation of US nodes and international archives

- Being realized through PDS4
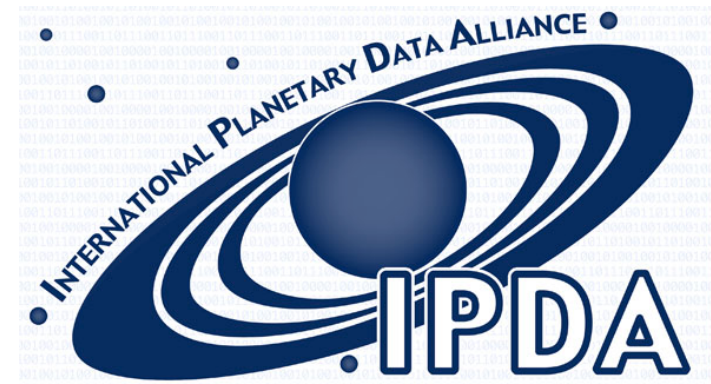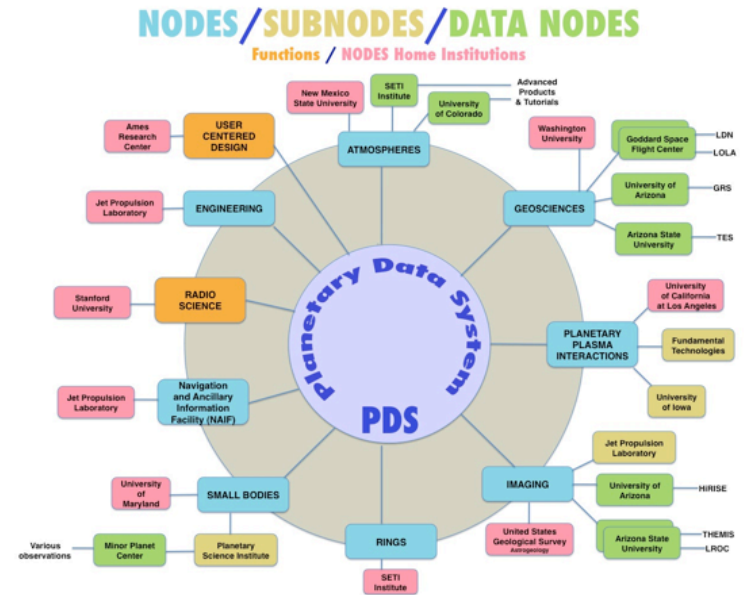
Key Drivers Impacting PDS

- More Data
- More Complexity (instruments, data)
- More Producer Interfaces
- Greater User Expectations
- Archive vs. Usability

- Funding Constraints
- Creating a system from the federation
- Internationalization
- Increasing IT security threats
- New Technology Opportunities

"Support the ongoing effort to evolve the Planetary Data System from an archiving facility to an effective online resource for the NASA and international communities." -- Planetary Science Decadal Survey, NRC, 2013-2022

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

U.S. Planetary Data Archives (TBs)
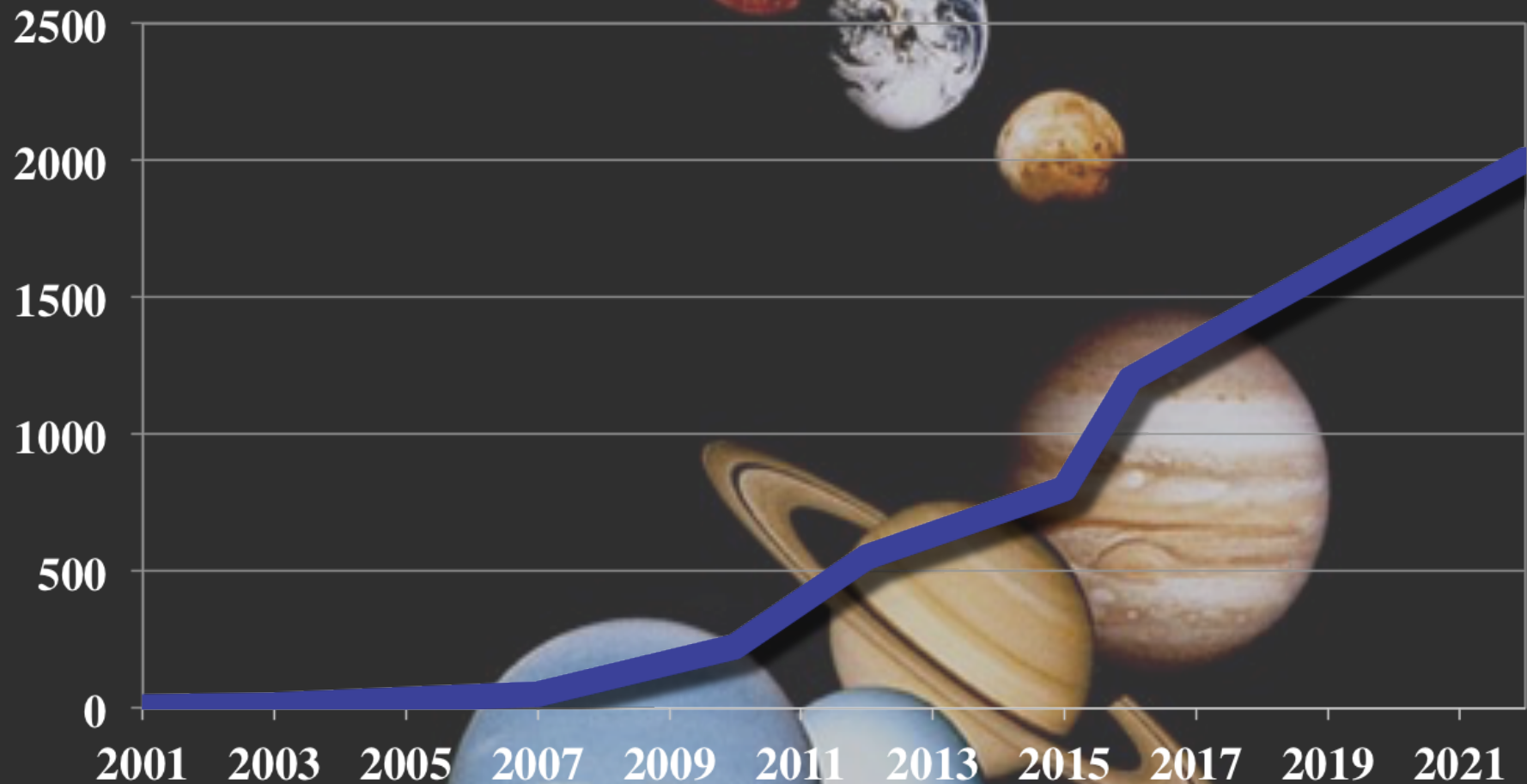
Diversity of the PDS

Type of Data	Distinct Products
Data Sets	2151
Instrument Hosts	199
Instruments	625
Targets	4231
Missions/Investigations	71
Volumes	5847

- Total volume is currently ~1PB
- Over 40M data products
- Some missions have few instruments but many data products, e.g., LADEE
- Variety of data is the challenge

PDS4

- An international, information model-driven data architecture for distributed planetary data archives

- An explicit information model
  - Explicitly describe the diversity of planetary data
  - Drive the definition of data to enable management, search and analytics across PDS and IPDA

- Distributed software services architecture
  - Services both within PDS and at international partners
  - Consistent protocols for access to the data and services
  - A distributed registry and search infrastructure
  - Tools that are built on top of the PDS4 information model

PDS4 is co-developed with the community

PDS4 Information Model

- PDS4 Information Model plays a key role in defining the data and its relationships
  - Defines explicit relationships between major entities of the PDS
  - Establishes an overarching governance model for PDS data
  - Handles the diversity of different disciplines
  - New instruments, observation types and data can be accommodated
  - PDS labels are tied to the model to increase consistency

- Changes managed through a change control board with members selected from PDS and IPDA

PDS4: A Model-Driven Strategy

Information System Architecture

Information Architecture

System Model
- Information Object
- Identification
- Referencing
- State

Domain Model (governance levels)
- Top Level
  - Representation/Format
  - Context, Provenance, Integrity
- Domain
  - Science
  - Engineering
  - Exploration
- Missions/Systems
  - Satellite/Airborne
  - Mission Operations

PDS4 Information Model

configure

System Architecture

describe

Configurable Components
- Data Management Model
- Search/Access Model
- Analytics Model

drive

Configured System

use

produce

Planetary Data

Crichton, D. Hughes, J.S. ; Hardman, S. ; Law, E. ; Beebe, R. ; Morgan, T.; Grayzeck, E.
A Scalable Planetary Science Information Architecture for Big Science Data.
IEEE 10th International Conference on e-Science, October 2014.

8

Core System Builds

- PDS4 uses system builds to bring together the software and the information model
  - Established very early in the project to organize releases
  - Provides a predictable structure to bring the teams together
  - Provides incremental functionality relative to budget constraints
  - *Support co-development with the community*

- Each build provides a full lifecycle to capture, CM, integrate, test and deploy the release

- V1.8 released in April 2017

Core Software and Standards Deliverables

Software System

- Registry Service
- Harvest Tool
- Validate Tool
- Security Service
- Report Service
- Search Service
- Transform Tool
- Catalog Tool
- Tools for different types of registries
- PDS4 Libraries
- Upgraded portal search and page views to support PDS4

Data Standards*

- Information Model
- XML Schemas
- Data Dictionary
- Concepts Document
- Standards Reference
- Data Providers Handbook
- PDS4 Example Products

* Posted to https://pds.nasa.gov/pds4

System Builds and Governance Approach

Discipline Node/IPDA Info Models → Distribute to Community

Core PDS System, Model, Tools, Libraries

Transfer to Core

PDS Management, Requirements →

(1) Information Models
(2) Software

PDS Core System Components

Change Control Board

System Builds (Software+ Model Integration, Regression Testing, etc)

System Release

→ Distribute to Community

Transfer to Core

Discipline Node and Community Software Tools → Distribute to Community

Major releases every 6 months

Support co-development with the community

Planetary Tools Registry

Tool Registry

This interface enables search and discovery of tools, services, and APIs for working with data following the PDS standards. Tools have been submitted from the broad PDS community and multiple institutions, including those from members of the International Planetary Data Alliance (IPDA). This interface allows the user to search for and discover these tools. The interface also allows tool providers to submit their software for inclusion in the registry.

Browse or search the PDS tool registry. Select a tool below to view the details.

Search

Category:
- All
- Analysis
- Reader
- Design
- Dissemination
- Generation
- Planning
- Search
- Transformation
- ● Validation
- Visualization

Interface Type:
- ● All
- API
- Command-Line
- GUI
- Service

Support:
- ● Both
- PDS3
- PDS4

Tools

Displaying 1 to 5 of (5) results.
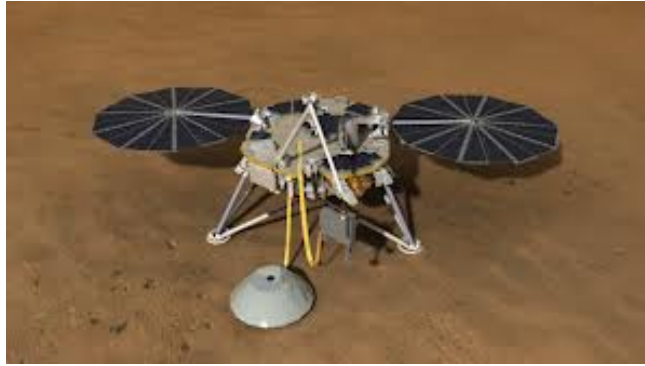
Name	PDS Version
Label Validation Tool (VTool)	PDS3
A Java-based command-line tool used for validating PDS3 data product labels.	
Online Peer Review Tool	PDS4
This is an on-line tool to homogenize review efforts for smaller datasets in either PDS3 or PDS4 label formats.	
PDS Tools Package	PDS3
The PDS Tools Package is the complete set of supported legacy tools for accessing and reading PDS3 data and labels.	
PSA Volume Verifier (PVV)	PDS3
The PVV allows for validation and delivery of a scientific dataset for ingestion to the Planetary Science Archive (PSA).	
Validate Tool	PDS4
A Java-based command-line tool used for validating PDS4 product labels and product data.	

Quick Searches
- Mars Science Laboratory
- Mercury
- Venus
- Mars
- Jupiter
- Saturn
- Uranus, Neptune, Pluto
- Rings
- Asteroids
- Comets
- Planetary Dust
- Earth's Moon
- Solar Wind

PDS Nodes
- Atmospheres
- Geosciences
- Cartography and Imaging Sciences
- Navigational & Ancillary Information (NAIF)
- Planetary Plasma Interactions (PPI)
- Ring-Moon Systems
- Small Bodies

PDS Support
- Management
- Engineering

https://pds.nasa.gov/tools/tool-registry

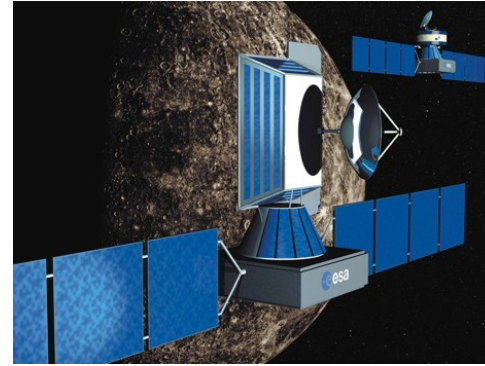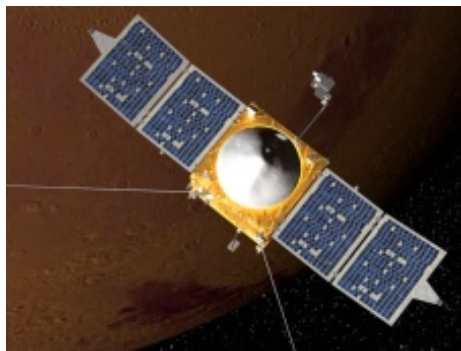# International Collaboration on PDS4 through the IPDA

LADEE (NASA)

InSight (NASA)

BepiColombo (ESA/JAXA)

MAVEN (NASA)
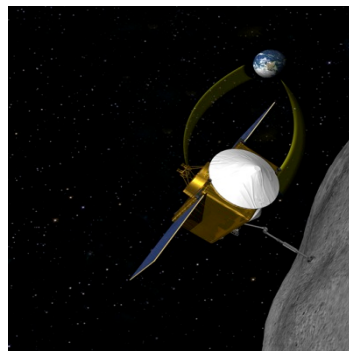
Osiris-REx (NASA)

ExoMars (ESA/Russia)

JUICE (ESA)

...also Hayabussa-2, Chandrayaan-2

Mars 2020 (NASA)

13

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

International Registration, Search, and Access

Search Service supports the PDS and PDAP protocols enabling development of other portals and applications on this infrastructure.

Registered Objects:
- Websites
- Data Sets
- Investigations, Instruments, etc.
- Tools and Services

Search results include mission support pages and other more specific search interfaces.

14
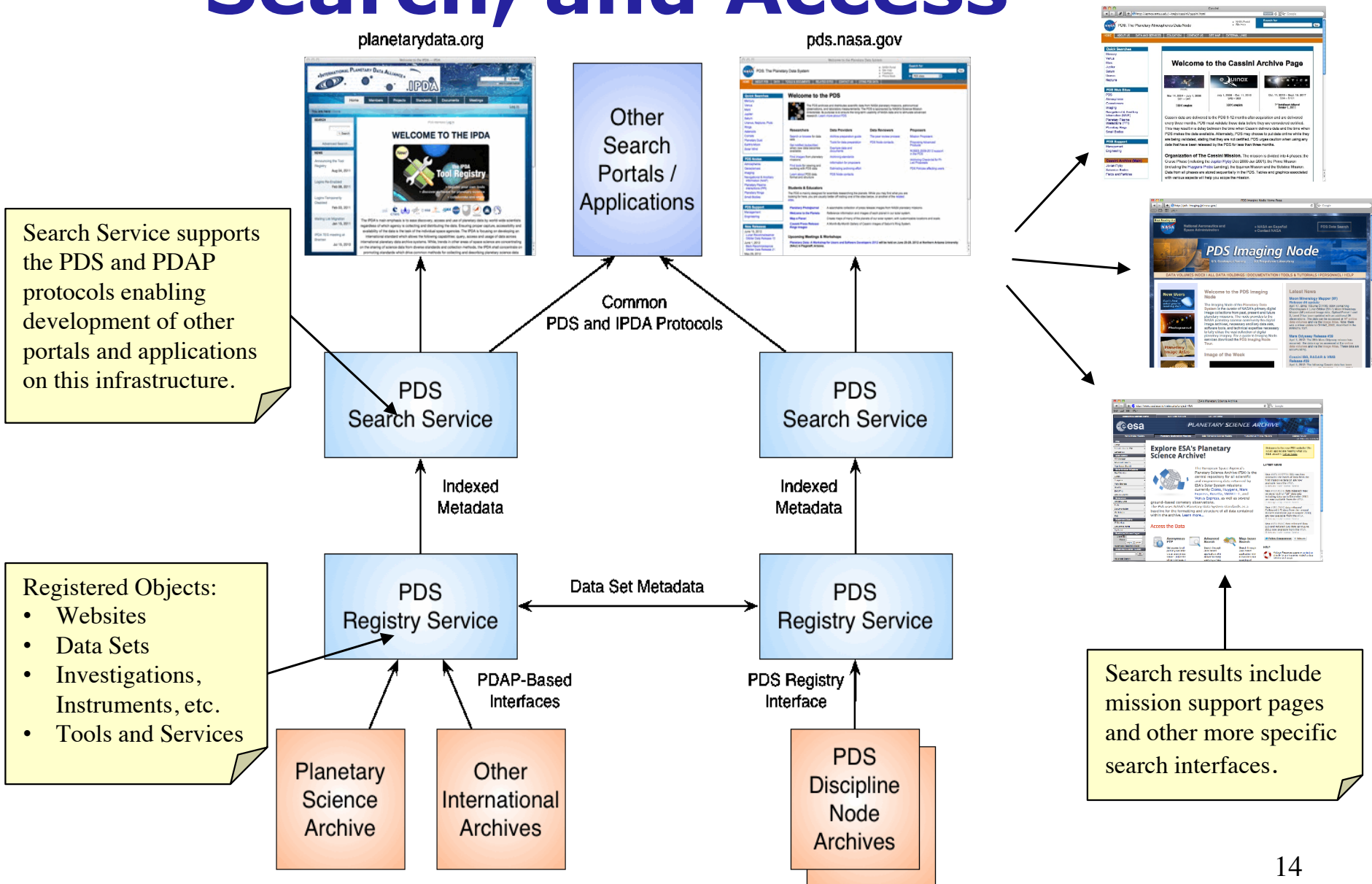
Future: An International Platform for Planetary Data Archiving, Management, and Research
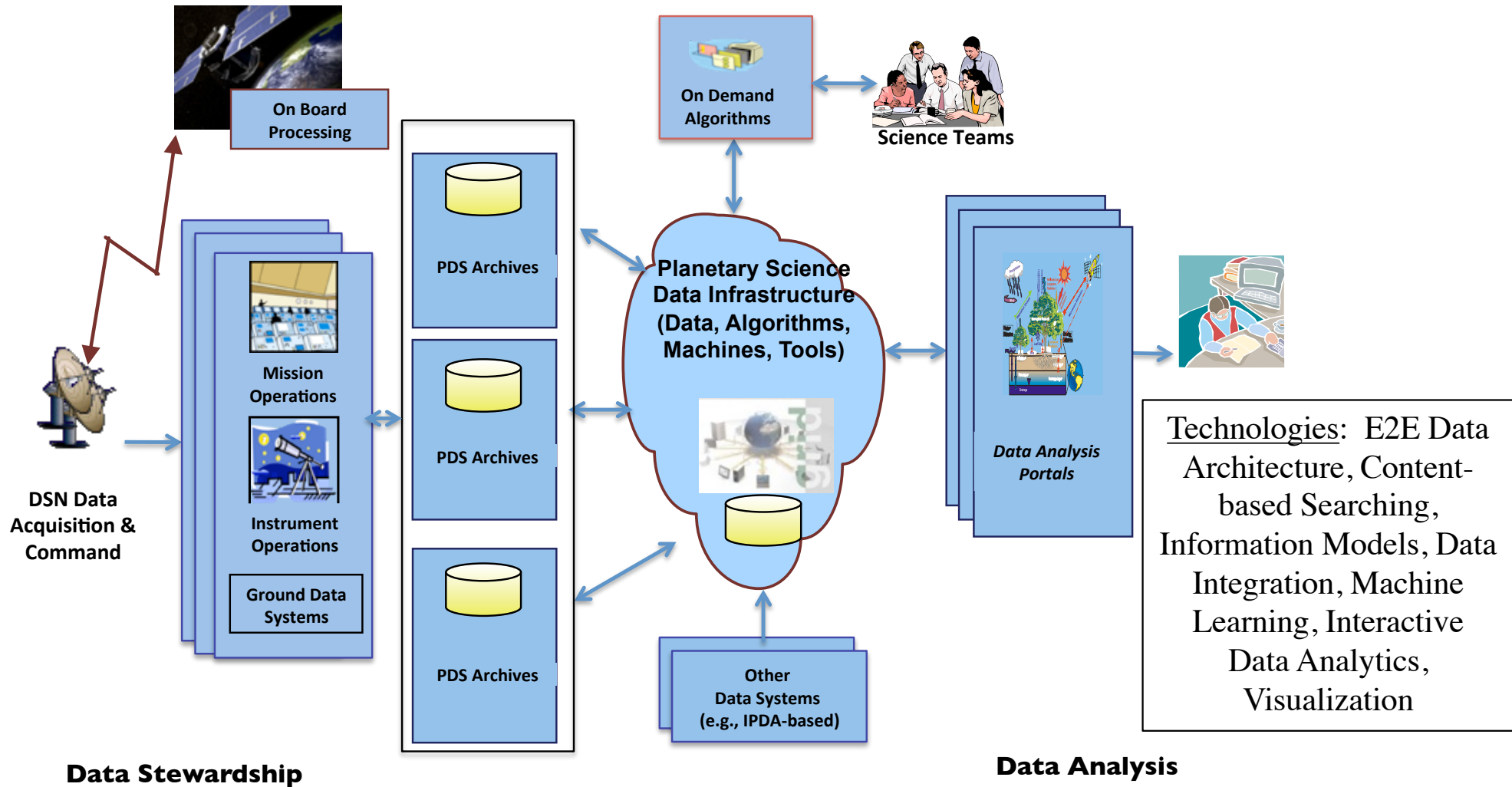
On Board Processing

On Demand Algorithms

Science Teams

DSN Data Acquisition & Command

Mission Operations

Instrument Operations

Ground Data Systems

PDS Archives

PDS Archives

PDS Archives

Planetary Science Data Infrastructure (Data, Algorithms, Machines, Tools)

Other Data Systems (e.g., IPDA-based)

Data Analysis Portals

Technologies: E2E Data Architecture, Content-based Searching, Information Models, Data Integration, Machine Learning, Interactive Data Analytics, Visualization

Data Stewardship

Data Analysis

"Support the ongoing effort to evolve the Planetary Data System from an archiving facility to an effective online resource for the NASA and international communities." -- Planetary Science Decadal Survey, NRC, 2013-2022

Summary

- PDS has now released eight versions of PDS4
  - Stable architecture, models, and software
  - Increasing capabilities over-time; governance enables it to be developed with the community

- PDS4 implemented or being implemented on a growing set of international missions

- Good progress towards international interoperability and data sharing

- Increasing opportunities to enable data-driven approaches
  - *See upcoming Planetary Data Analytics Workshop in April 2018*

Backup

Considering Future Data Intensive Approaches to Data Lifecycle Model
for NASA Planetary Missions

Emerging Solutions
- *Onboard Data Products*
- *Onboard Data Prioritization*
- *Flight Computing*
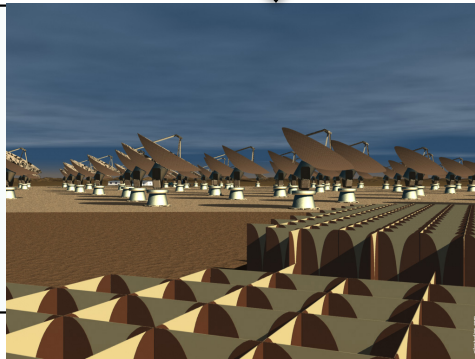
(1) Too much data, too fast; cannot transport data efficiently enough to store

Observational Platforms /Flight Computing

Massive Data Archives and Big Data Analytics

Emerging Solutions
- *Low-Power Digital Signal Processing*
- *Data Triage*
- *Exa-scale Computing*

Emerging Solutions
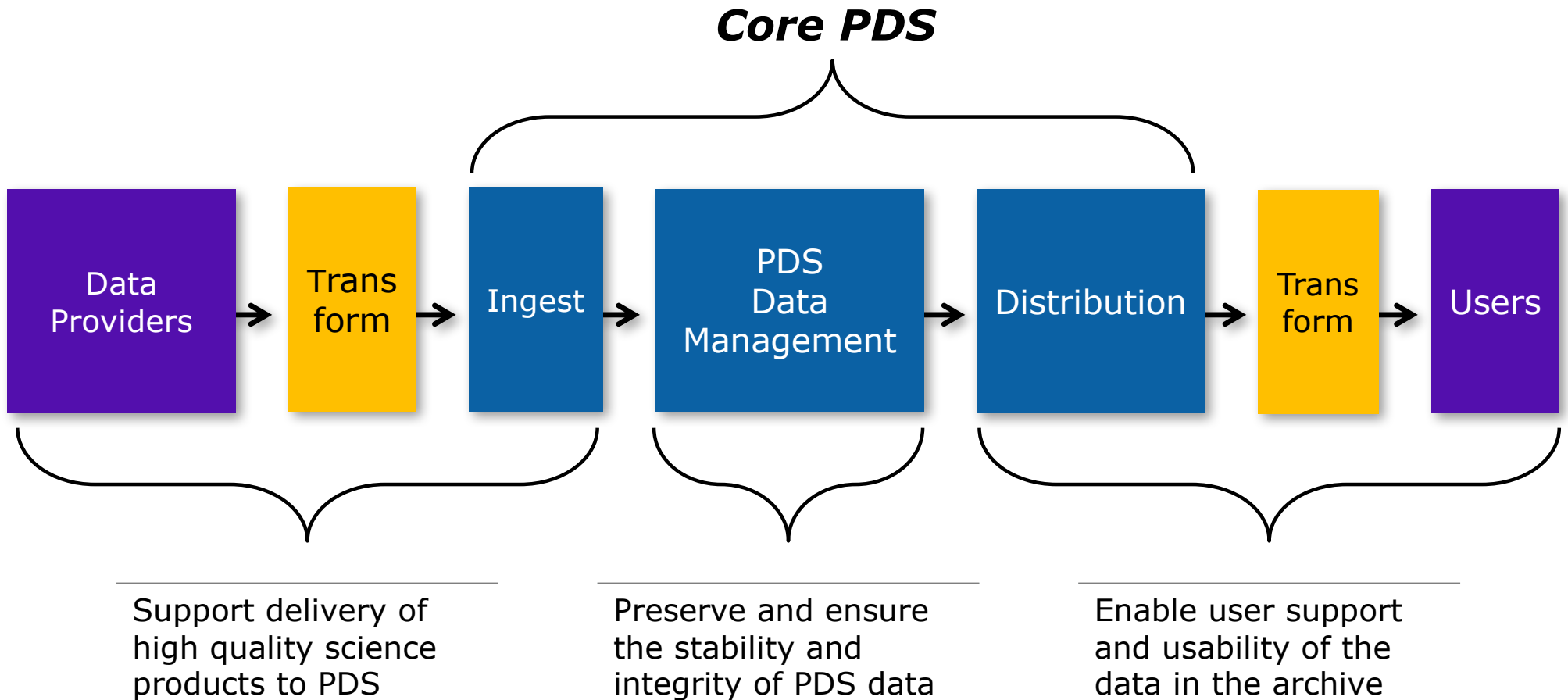- *Distributed Data Analytics*
- *Advanced Data Science Methods*
- *Scalable Computation and Storage*

© 3poD ® www.ClipartOf.com/15304

(2) Data collection capacity at the instrument continually outstrips data transport (downlink) capacity

(3) Data distributed in massive archives; many different types of measurements and observations

Ground-based Mission Systems

18

Major PDS Functions

Core PDS

Data Providers → Trans form → Ingest → PDS Data Management → Distribution → Trans form → Users

Support delivery of high quality science products to PDS

Preserve and ensure the stability and integrity of PDS data

Enable user support and usability of the data in the archive

19

PDS Level 1 Requirements

1. 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

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

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

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