

PDS4 Update

Dan Crichton

August 2014

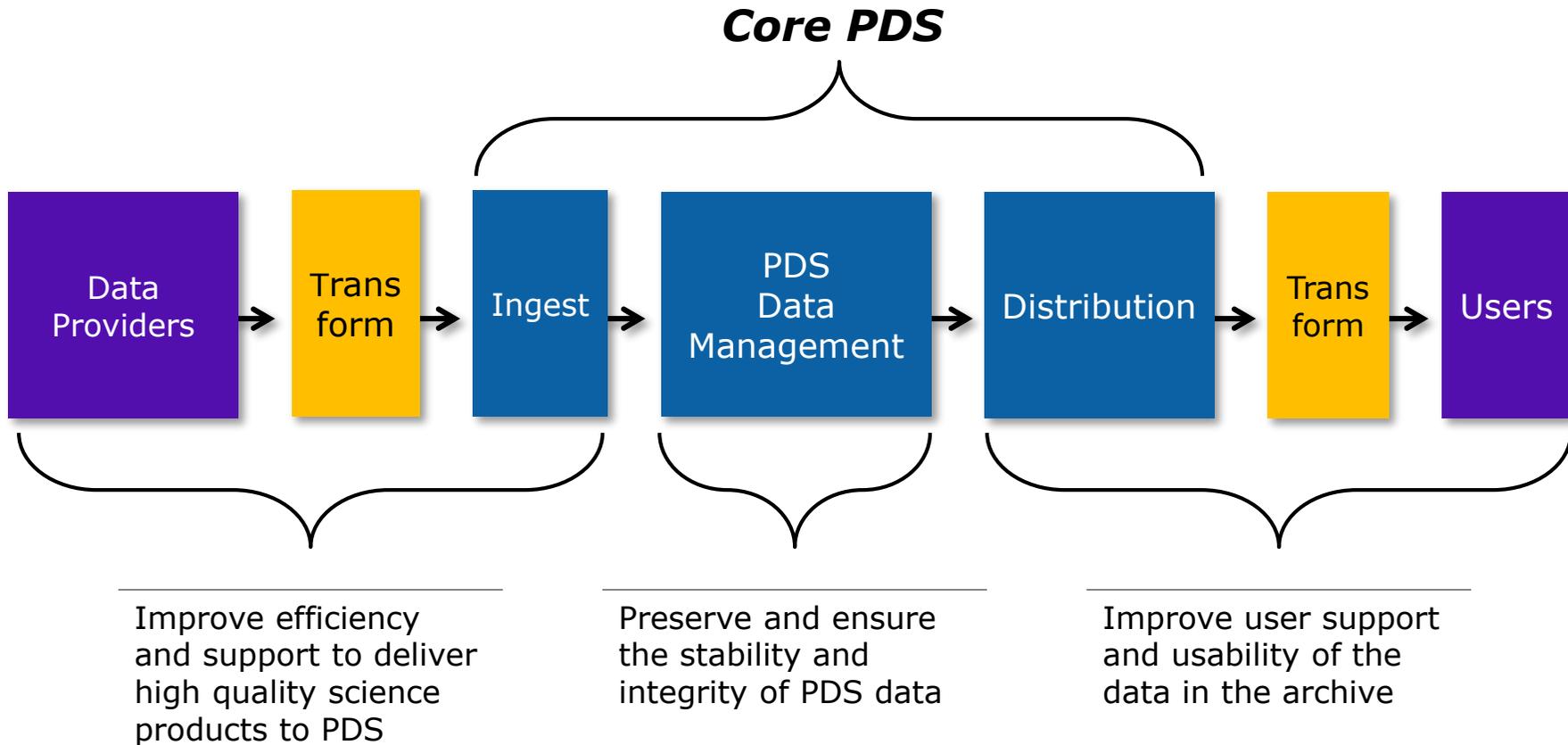
PDS4 MC Topics

- IPDA Report – Dan Crichton/Tom Stein
- PDS4 Report – Dan Crichton
- CCB –Lynn Neakrase
- IM/DDWG – Steve Hughes
- Software – Sean Hardman
- Tools – All

PDS4: The Next Generation PDS

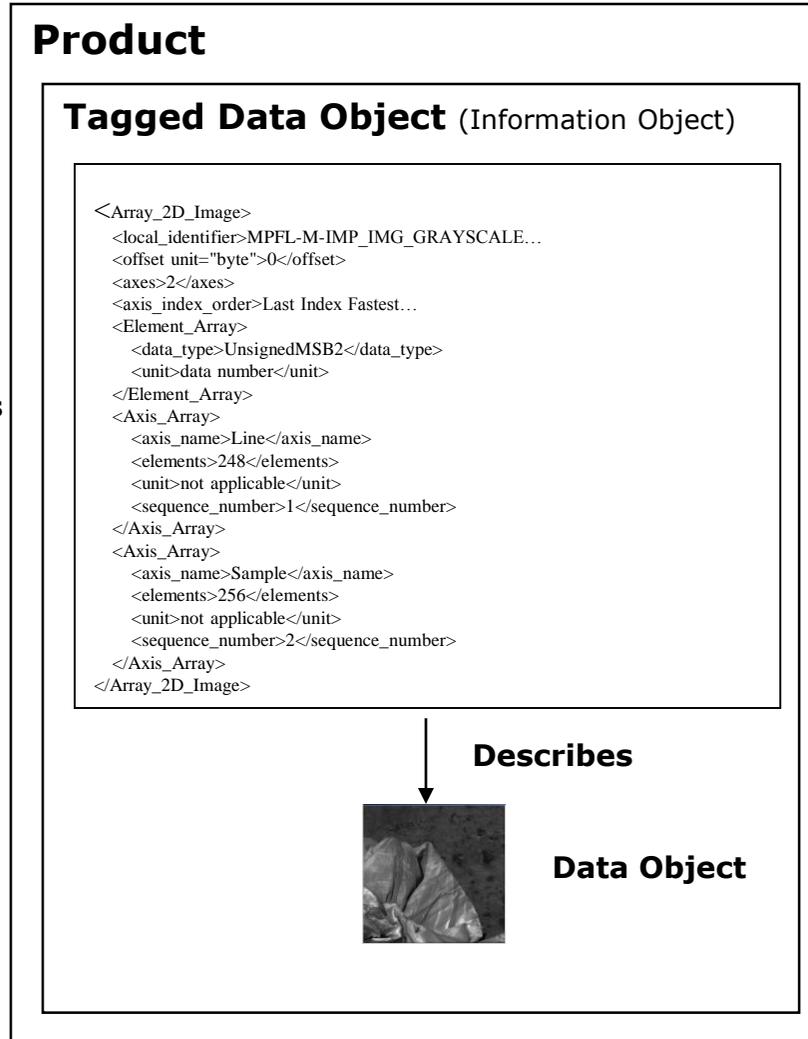
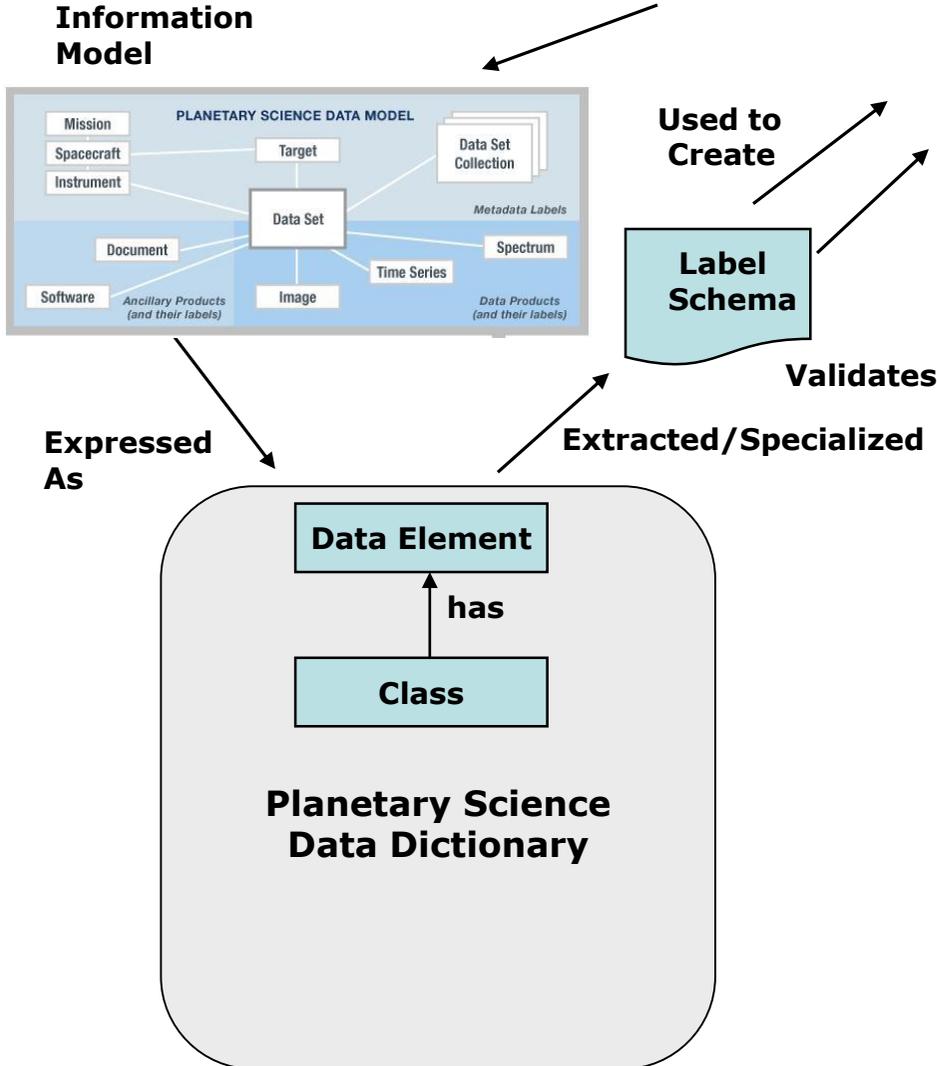
- PDS4 is a PDS-wide project to upgrade from PDS version 3 to version 4 to address many of these challenges
- An **explicit information architecture**
 - All PDS data tied to a common model to improve validation and discovery
 - Use of XML, a well-supported international standard, for data product labeling, validation, and searching.
 - A hierarchy of data dictionaries built to the ISO 11179 standard, designed to increase flexibility, enable complex searches, and make it easier to share data internationally.
- An **explicit software/technical architecture**
 - 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 in PDS
 - A distributed search infrastructure

Challenge: End-to-End System and Data Integration



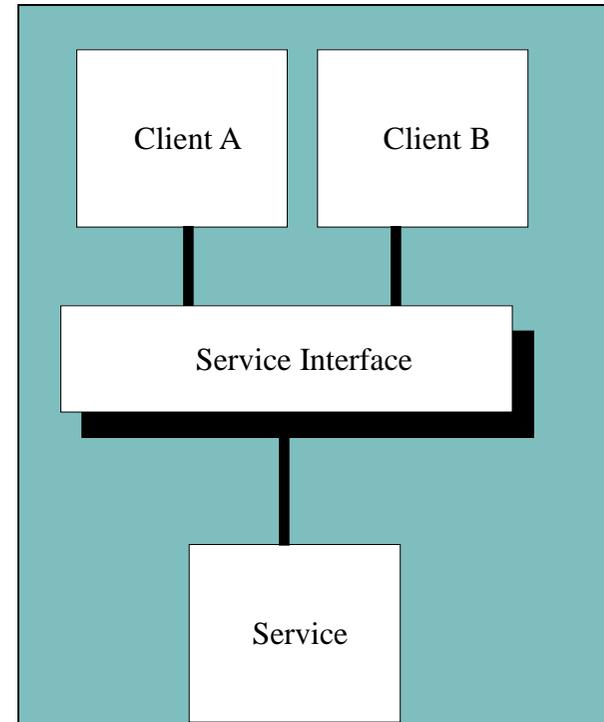
Information Architecture Concepts

Design/change starts here



System Design Approach

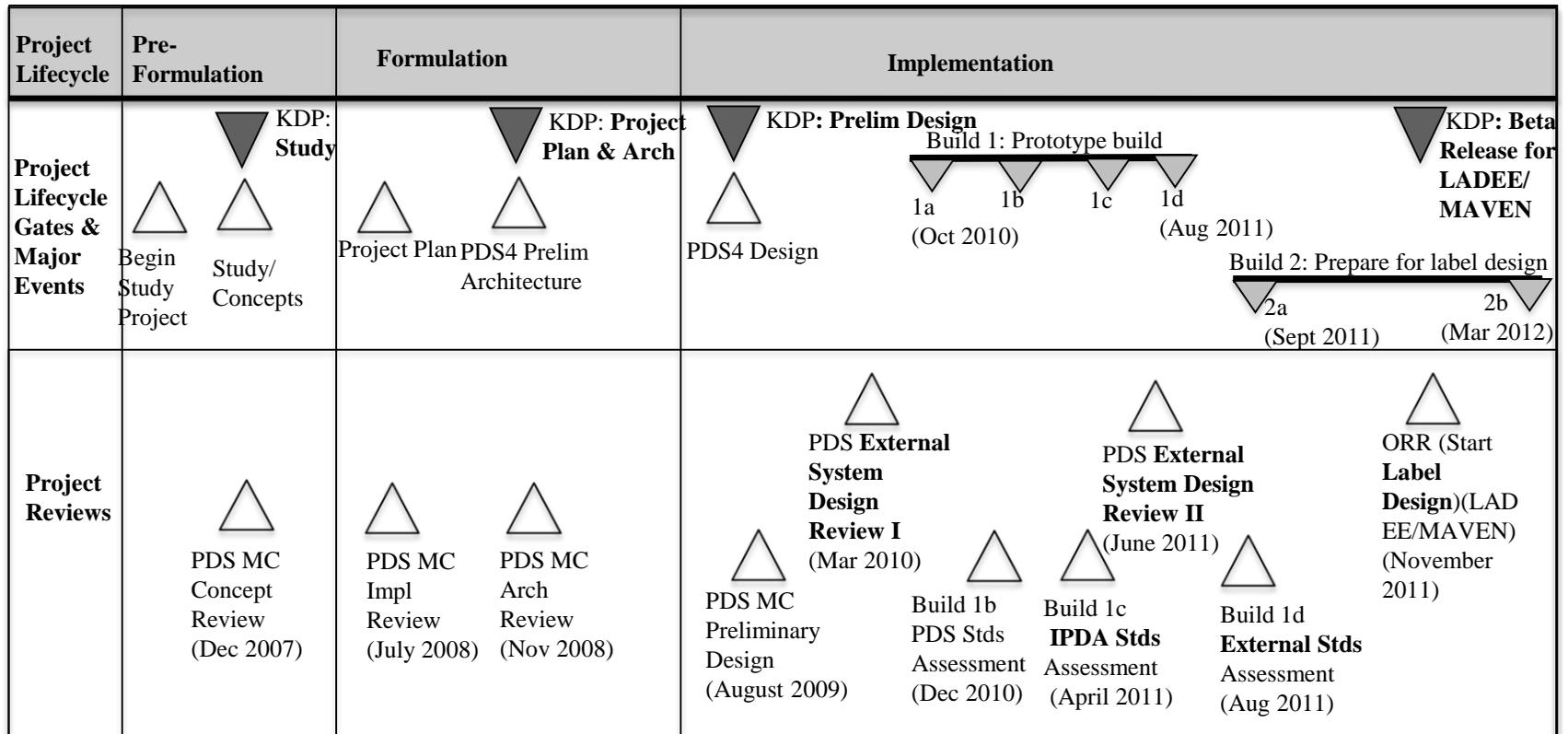
- Based on a distributed information services architecture (aka SOA-style)
 - Allow for common and node specific network-based (e.g., REST) services.
 - Allow for integrating with other systems through IPDA standards.
- System includes services, tools and applications
- Use of online registries across the PDS to track and share information about PDS holdings
- Implement distributed services that bring PDS forward into the online era of running a national data system
 - With good data standards, they become critical to ultimately improving the usability of PDS
 - Support on-demand **transformation** to/from PDS



Summary of Progress to Date

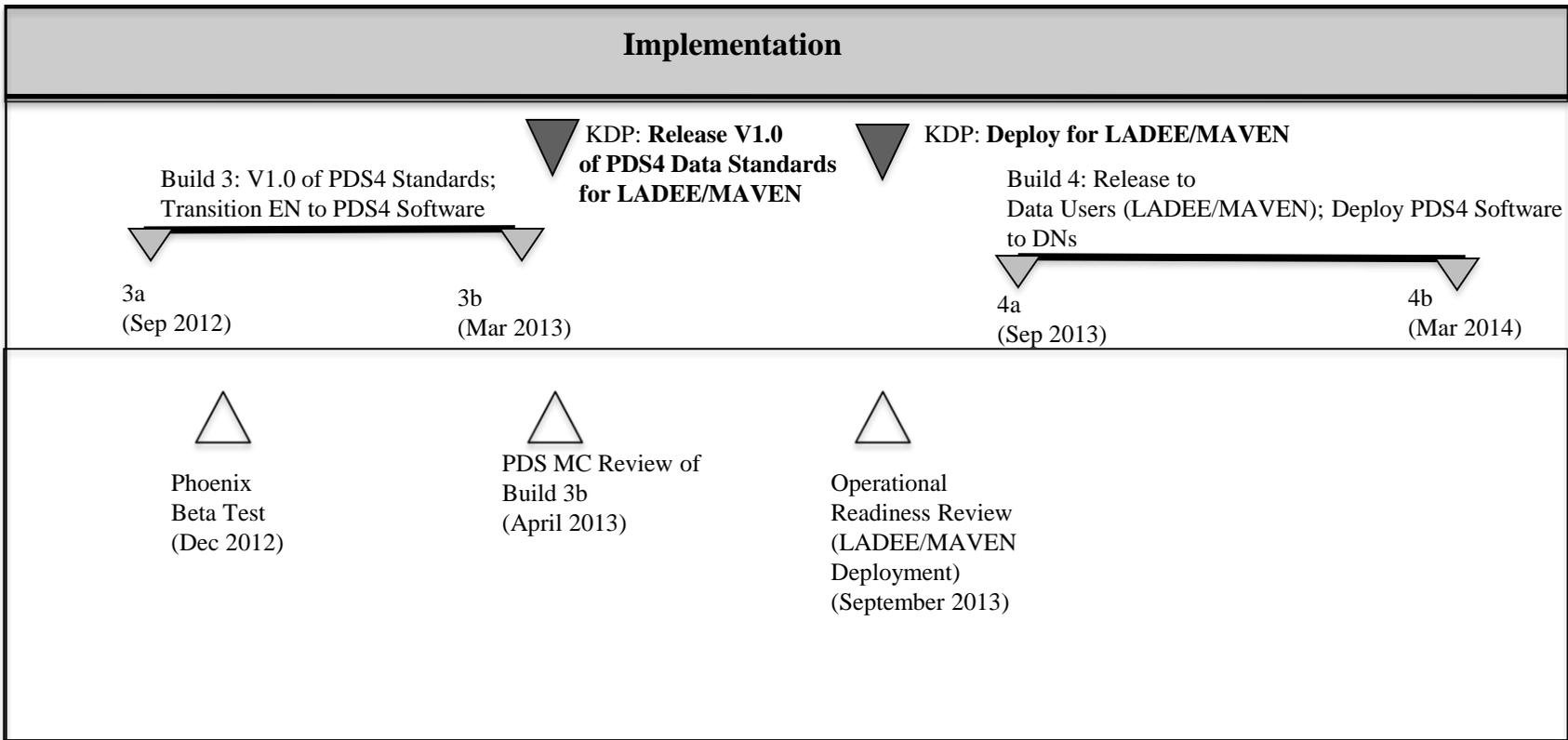
- Initial requirements in place
- PDS-wide system architecture defined
- Major reviews conducted (Design Review 1 and 2, ORR)
- System builds grouped by purpose: build 1,2,3 and 4
 - Iteratively increase capability and stability
- Operational capabilities deployed
 - EN fully running PDS4 software supporting access to both PDS3 and PDS4 services at nodes mitigating migration pressure
- Change control board established
 - JIRA deployed to manage tracking
- Product development underway at nodes and internationally
 - Initial peer reviews conducted
- First mission getting ready to do data distribution (LADEE)
- IPDA endorsement of PDS4

Project Lifecycle thru Build 3

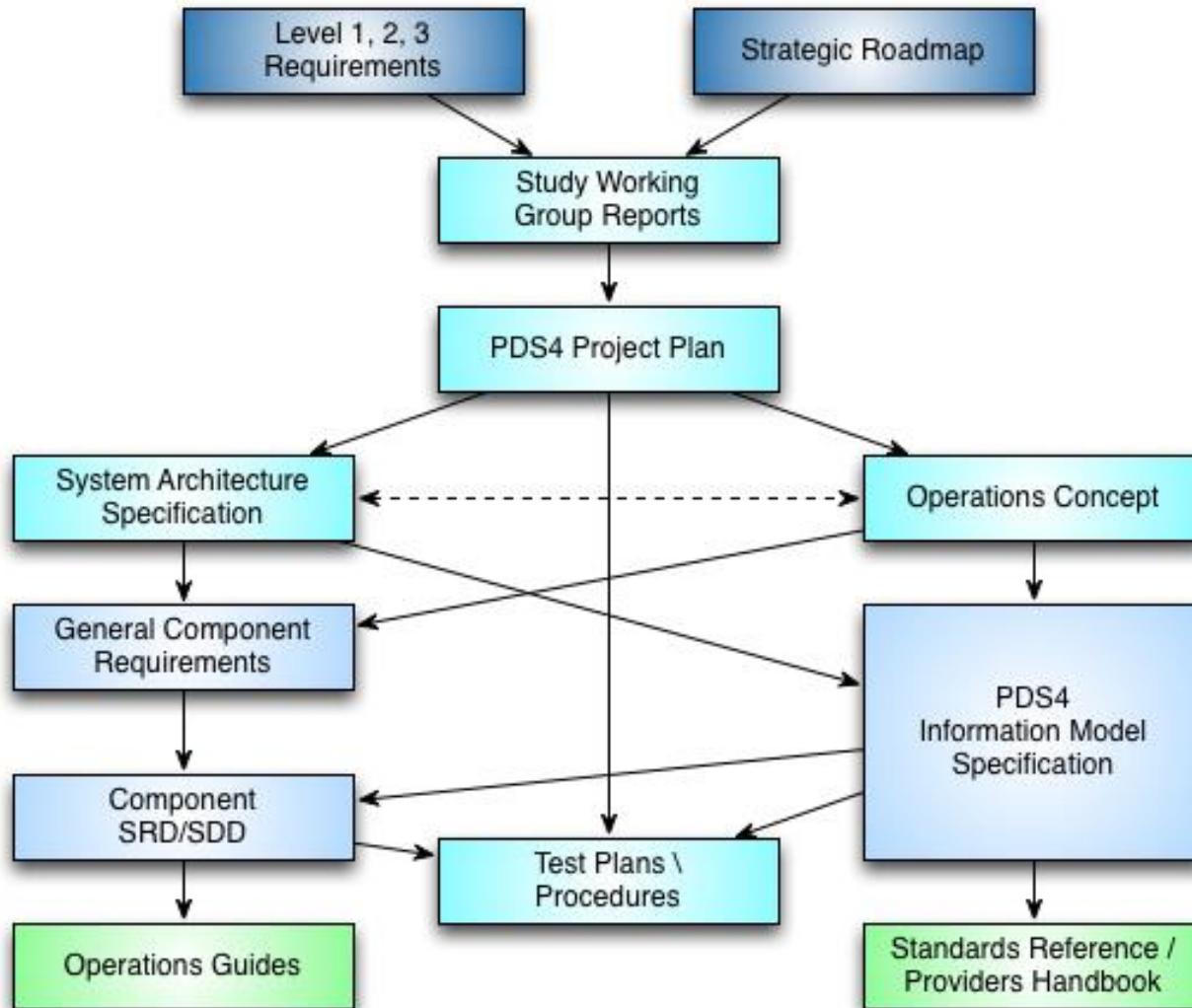


Architecture, requirements, design, test, releases posted at:
<http://pds-engineering.jpl.nasa.gov>

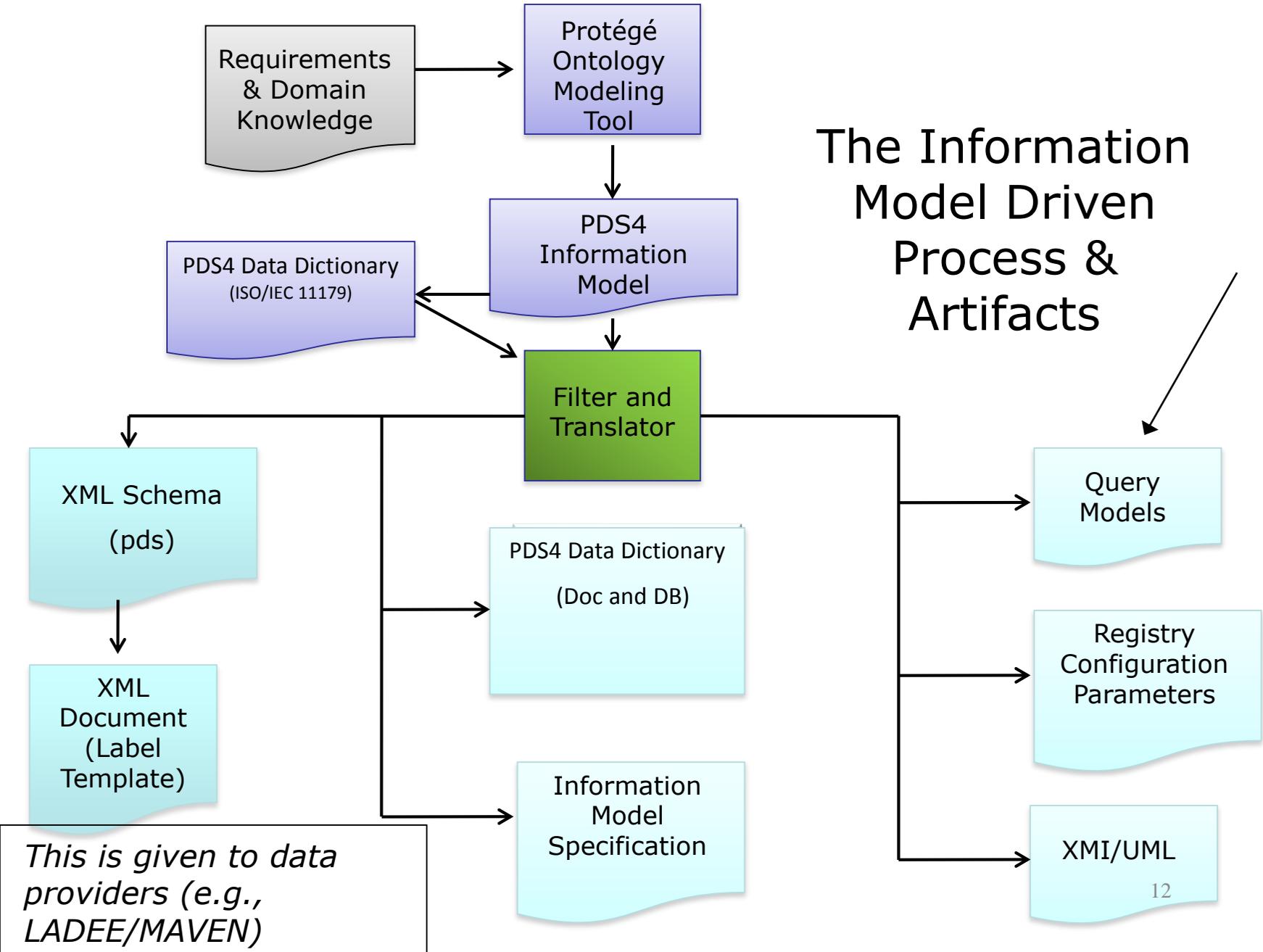
Build 4+



Document Tree



The Information Model Driven Process & Artifacts



This is given to data providers (e.g., LADEE/MAVEN)

PDS4 Builds

- Build 3b (April 2013)
 - V1.0 Standards derived from build 3b
 - V1.0 classes under strict CM
 - Scoped to LADEE/MAVEN product design needs
- Establish CCB for V1.0 of PDS4
- Build 4a (September 2013)
 - V1.1 Standards including additional classes as discussed at PDS MC
 - Initial User Services
- ORR: Distribution capabilities and plans for LADEE, MAVEN (September 2013)
- Build 4b (March 2014)
 - Additional user services/improvements
 - Ready for LADEE
 - Support for InSight/SEIS SEED data files
- Now getting ready for build 5a to begin I&T at end of September
 - November MC can be used to discuss I&T results and deployment

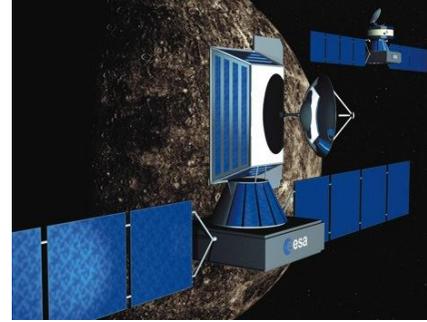
PDS4 Planned Mission Support



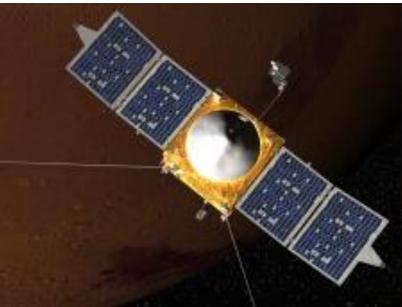
LADEE (NASA)



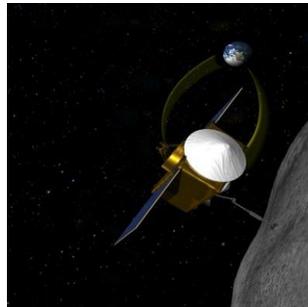
InSight (NASA)



BepiColumbo (ESA/JAXA)



MAVEN (NASA)



Osiris-Rex (NASA)



**ExoMars
(ESA/Russia)**

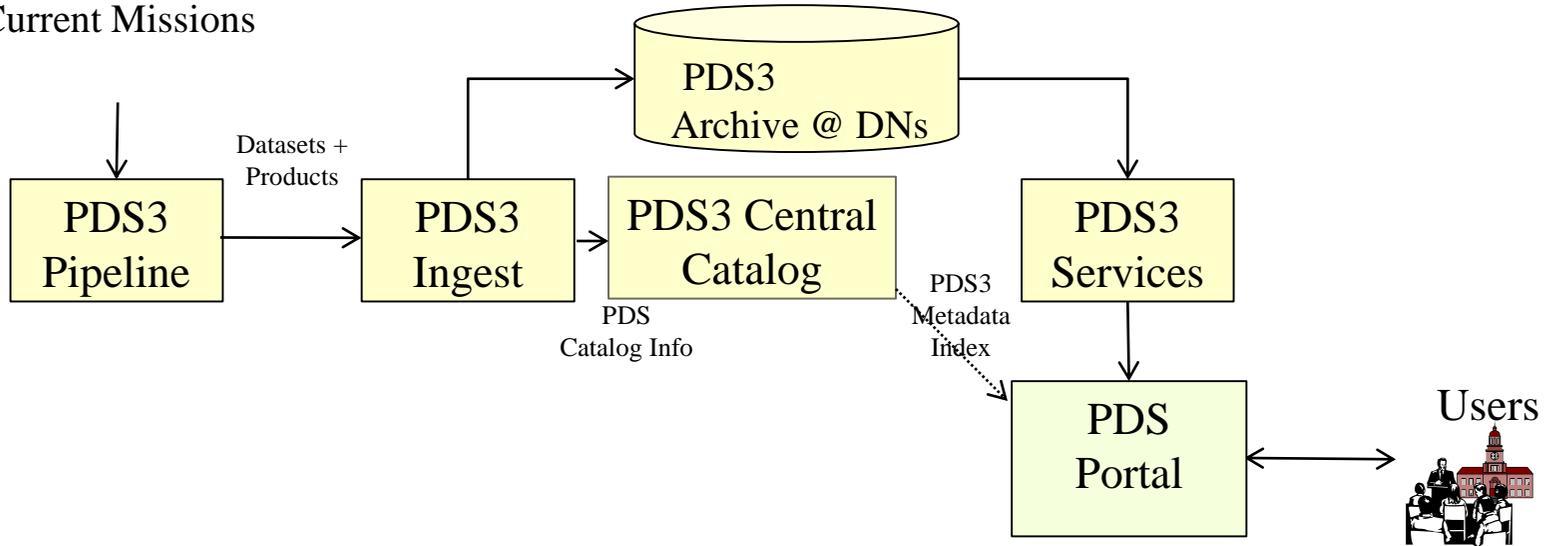


**JUICE
(ESA)**

...also Hyabussa-2, Chandryaan-2

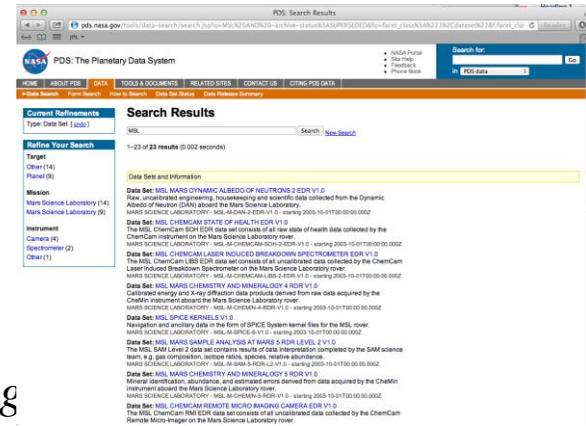
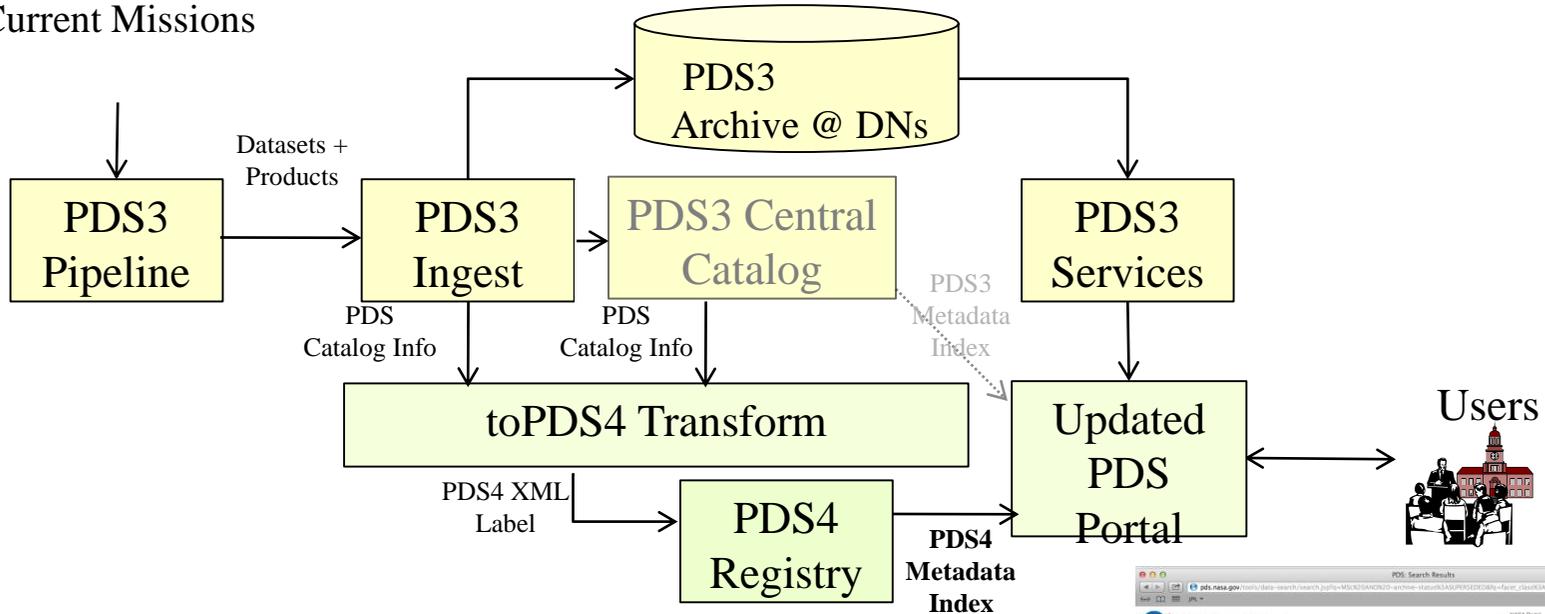
PDS3 Implementation

Current Missions



Transition

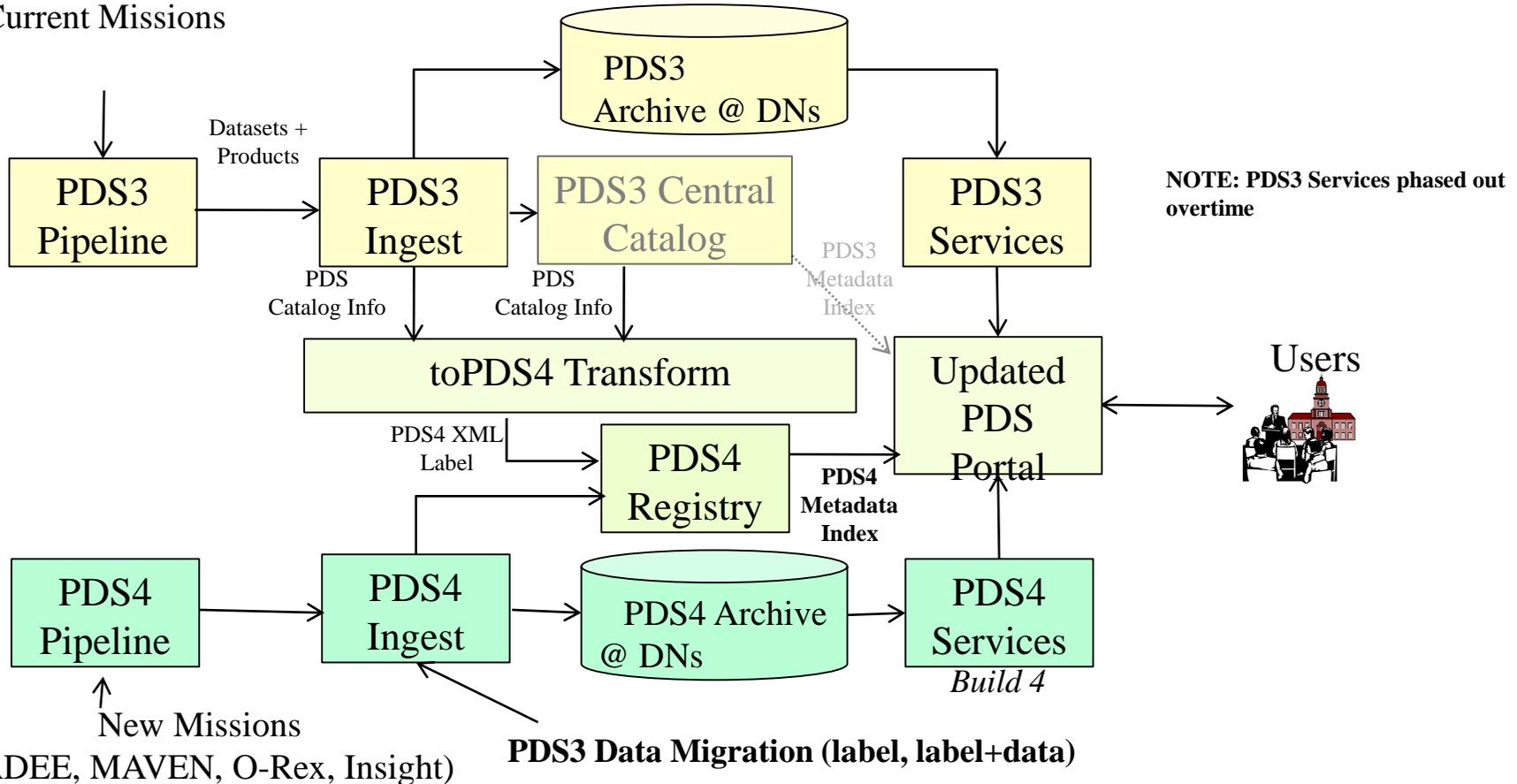
Current Missions



(1) *PDS4 infrastructure deployed at EN; Central catalog migrated.*

Support for LADEE/MAVEN

Current Missions



(1) *PDS4 infrastructure deployed at EN; Central catalog migrated.*

(2) *Minimized migration pressure.*

(3) *Working towards acceptance/distribution of new PDS4 mission data*

PDS4 Policies

- The following PDS4 specific policies are posted to <http://pds.nasa.gov/policy>
 - PDS Policy on Formats for PDS4 Data and Documentation (June 2014)
 - PDS Policy on Data Processing Levels (March 2013)
 - PDS Policy on Transition from PDS3 to PDS4 (November 2010)

Proposed Validation Policy

- Improve the quality of PDS4 bundles from data providers by ensuring that supplied tools coupled with manual verification is required prior to delivery of data to a node (or international partner)
- Discussed at April F2F and IPDA F2F
- MC needs to determine how to move forward
 - International considerations (full support for a policy + tool)
 - Tool considerations – agree to require a common tool
 - Document considerations (minor additions to DM Plans and PDS4 Standards Reference,

Proposed Validation Policy

Data providers delivering bundles to the PDS shall adhere to PDS4 standards by ensuring that the following criteria are met prior to delivery to PDS:

- 1) Syntactic validation: a) the XML label is validated against the appropriate schema rules; b) a mission schematron is syntactically correct; and c) a mission schema is syntactically correct.
- 2) Semantic validation: the XML label is validated against the appropriate schematron rules.
- 3) Content validation: the XML label accurately describes the data product.
- 4) Referential integrity: the relationships described, in and between digital objects described in the XML label, are consistent and represented

PDS supplied software validation tools support syntactic, semantic, specific content rules and referential integrity validation. Data providers must run these prior to delivery. Data Providers should use visual inspection to validate content that cannot be done programmatically (i.e., by using software validation tools).

Information Model Development

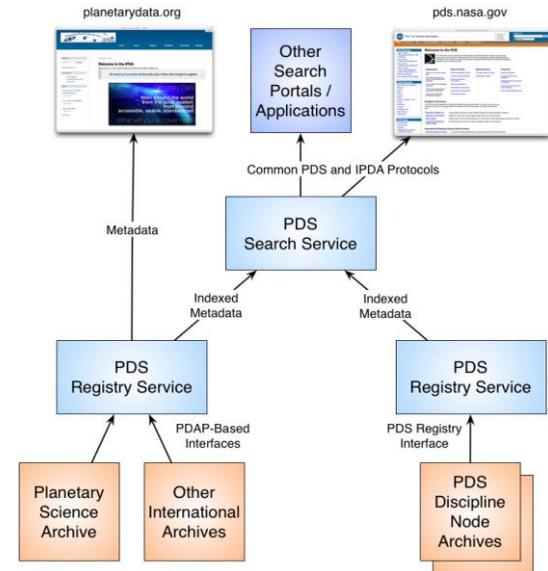
- Not a lot of substantial changes to the common model
 - Improvements and maintenance from experience
 - This is under CCB; Lynn will report out
- Discipline extensions being worked (e.g., cartography, geometry)
 - Discussion needed on how to best coordinate these (ESA request)
- Search and access discussions
 - Need to take advantage of modern search engine approaches across the PDS

Software and Tools Development

- Registry and search infrastructure has been running for a while now at EN
- Sean will discuss node deployments, but goal is
 - PDS3 tracked through a PDS3 registry (minimal label information) <- inconsistency in PDS3 makes this a challenge but we should begin tackling now
 - New PDS4 products <- this is model-driven
- Common tool development and releases occurring
 - Need to determine gaps (e.g., PDS4 version of NASAView, transformations, etc)
 - Will review the MC spreadsheet from April 2013 to ensure coordinated development

Search and Access

- The *search service* is based on the Apache Solr search engine
 - Fully open source
 - Can be customized for our planetary science model (and their disciplines)
 - **Can include multiple inputs for search (labels + other sources)**
 - Ranking is under our control
- Fully deployed at Engineering
 - Index includes data from multiple registries
- Adoption of a modern search engine infrastructure across PDS will allow us to tune overtime.



The screenshot shows the PDS website search results page. The header includes the NASA logo and the text 'PDS: The Planetary Data System'. A search bar at the top right contains the text 'cassini' and a 'Go' button. Below the search bar is a navigation menu with links for HOME, ABOUT PDS, PDS4, DATA, TOOLS & DOCUMENTS, RELATED SITES, CONTACT US, CITING PDS DATA, and POLICIES. The main content area is divided into two columns. The left column, titled 'Refine Your Search', lists various filters such as Agency (NASA (221), ESA (9)), Type (Data Set (187), Instrument (26), Target (7), Investigation (6), Instrument Host (4), Search Tool (2), Resource (1)), Target (Other (100), Planet (99), Satellite (65), Ring (22), Calibration (16)), and Investigation (Cassini-Huygens (179), Galileo (8), Voyager (4), Dawn Mission To Vesta And Ceres (1), Gravity Recovery And Interior Laboratory (1)). The right column, titled 'Search Results', shows the search results for 'cassini'. It indicates '1-50 of 233 results (0.002 seconds)'. The first result is 'Archive Information', which provides detailed information about the matching investigations and data sets. Below this is a 'Search Tools' section, which lists 'Cassini Image Search' and 'Outer Planets Unified Search (OPUS)'. The 'Data Sets and Information' section at the bottom lists 'Target: MASURSKY' and 'Target: FOMALHAUT'.

Build 5a

- Following the lifecycle we have established, we are planning for the next release
 - Need final CCB changes to the IM (V1.3)
 - Allow time for nodes to review changes prior to lock down for I&T
 - Will accept any node test products for our regression tests during I&T
 - I&T will exercise tests that integrate tools, services and data products
- Plan is to begin I&T October 1

Summary

- Policies
 - Need MC disposition on moving forward with a validation policy
- Build 5a planning is underway
 - Need node input
- Software
 - Need to begin registry population
 - Tools – core tools are emerging, but we need to work the gaps in our plan for FY15; will review spreadsheet
 - Transformations – we can continue to plan these
- User Services/Search Access
 - Move to a modern search engine infrastructure (e.g., search service)

Backup

Schedule

- https://pds-engineering.jpl.nasa.gov/pds2010/pds4_project_schedule.pdf