National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Enabling Virtual Observatory Access to Planetary Resources through PDS4

AGU 2015 - IN33E-02
16 December 2015

Sean Hardman[1], John Hughes[1], Daniel Crichton[1],
Baptiste Cecconi[2], Isa Barbarisi[3], Christophe Arviset[3]

[1] Jet Propulsion Laboratory, California Institute of Technology
[2] Paris Observatory Meudon
[3] European Space Agency

Topics

- Purpose and Background
- Overview of the PDS4
- Overview of the IVOA and EPN-TAP
- Proposed Testbed

Purpose

- Make planetary science resources available to the international space science community
  - *Develop an interface between the Planetary Data System (PDS) and the Virtual Observatory (VO) through the International Virtual Observatory Alliance (IVOA)*
  - *Build on previous work by the International Planetary Data Alliance (IPDA) to integrate international planetary data archives via common search protocols.*

Background

- PDS4, the next generation Planetary Data System (PDS), was developed using architectural principles that enable access to planetary science resources.
  - *Essentially all Solar System Exploration Science Data*
  - *Services, Tools, Documentation, and Expertise*

- EPN-TAP (Europlanet-VO Table Access Protocol) provides a good use case to demonstrate how an additional international protocol can be applied to allow access to international planetary science resources.

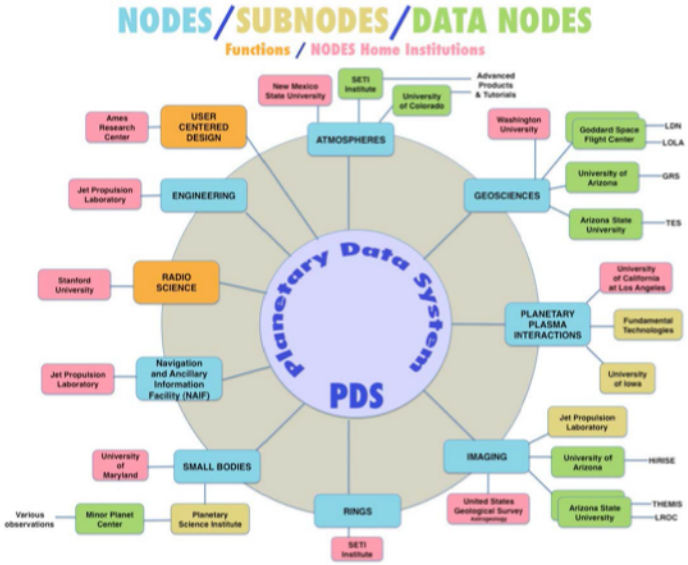
National Aeronautics and Space Administration

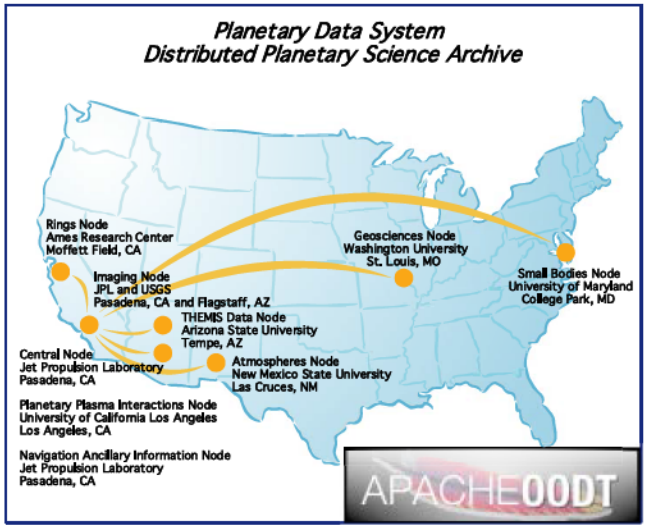
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Planetary Data System

- Purpose: 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.

- Infrastructure: a highly distributed infrastructure with planetary science data repositories implemented at major government labs and academic institutions
  - System driven by a well defined planetary science ontology
  - Approximately 900 TBs of data
  - Movement towards international interoperability
  - Implemented Apache OODT (developed @ NASA) in 2002 to share data; scaled to a SOA implemented in 2010

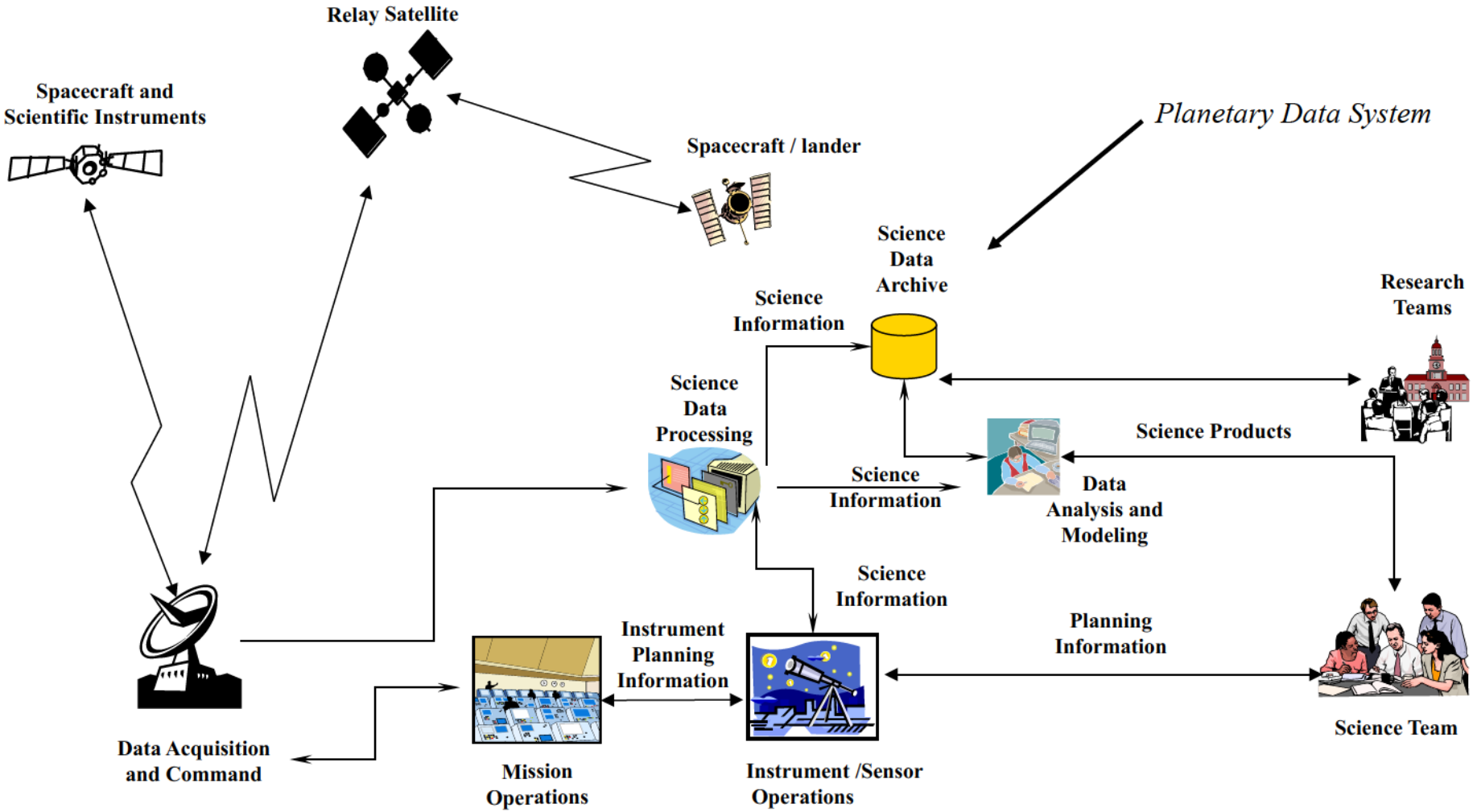




National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Reference Architecture Space Information Systems



Credit: CCSDS Reference Architecture for Space Information Management

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

PDS4: The Next Generation PDS

- The NASA Planetary Data System (PDS) after about 20 years of operations has developed PDS4, a major revision and transition to a modern system based on best practices for data system development.
- 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*
- An international collaboration for standards coordination and archiving and sharing data

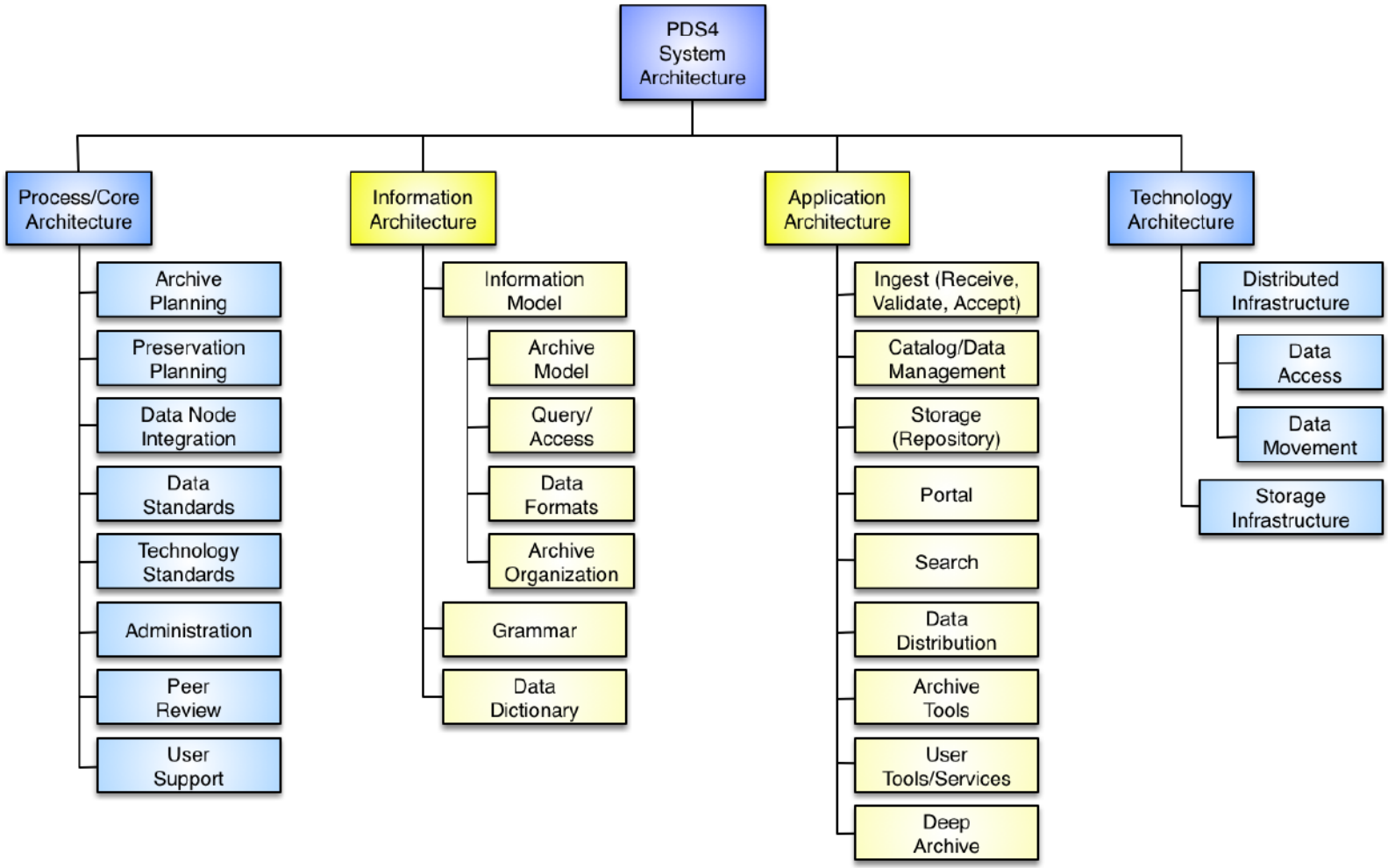
Key PDS4 Features

- Data Reuse
  - *Designed a few simple formats for 80% of the data*
  - *All things are formally defined once*
  - *Everything is registered as a product*
  - *Multi-level governance*
- Model Driven
  - *Model evolves with changes in the science discipline*
  - *Implementation technologies evolve at their own speed.*
  - *Improves interoperability at the information level across international data systems*
- Subsumes legacy archive
  - *Proxy labels exist for each legacy product*
  - *High value data sets are migrated as needed*

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

PDS4: System Architecture



Based on PDS Level 1, 2 and 3 requirements, the system architecture was decomposed into elements and organized accordingly.

PDS4 Information Architecture

- The information architecture is developed and maintained independent of the systems architecture

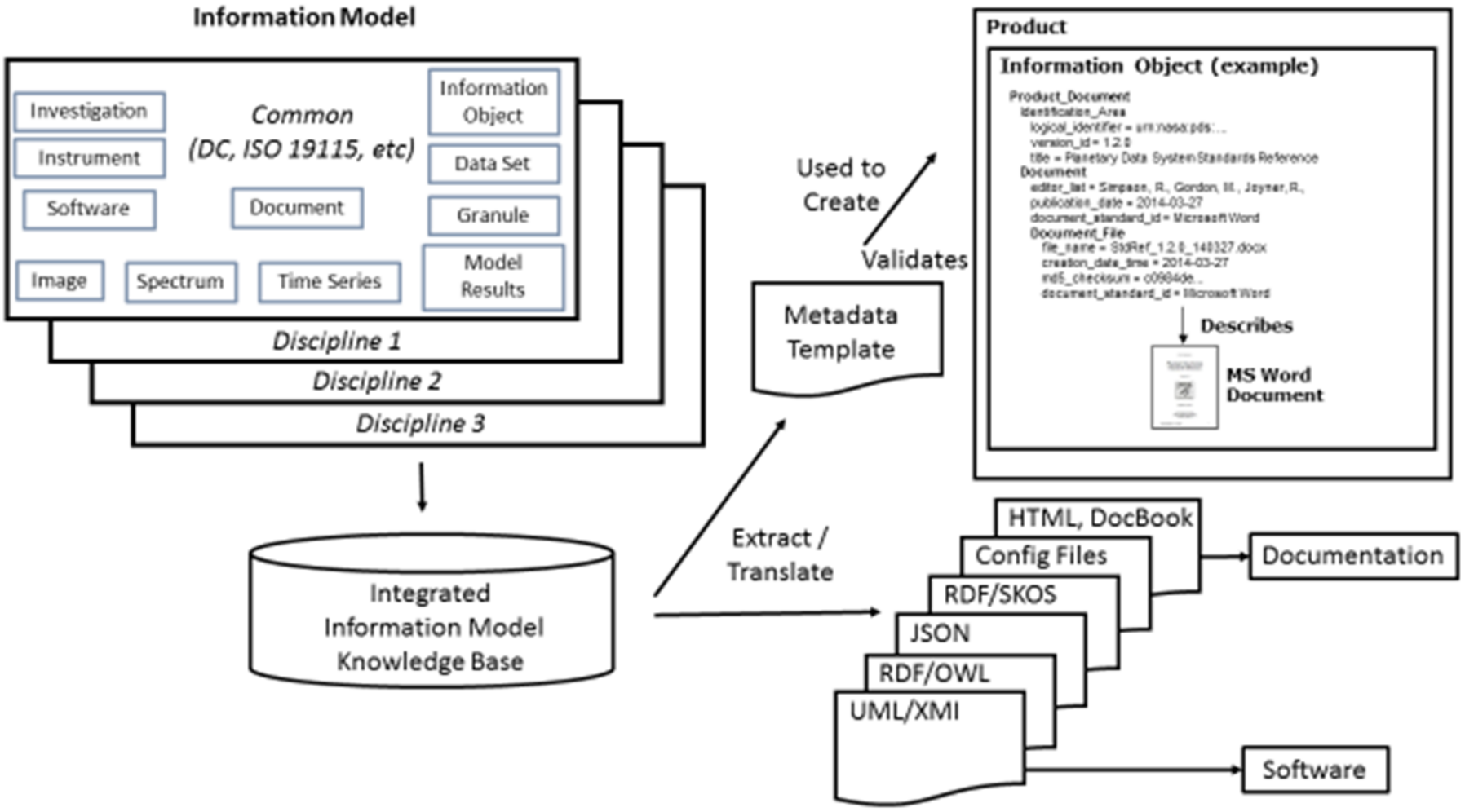
- Provides a formal, sharable, and stable set of requirements
  - *that enable understanding of the system*
  - *the configuration of system components*
  - *the basis for mapping to and from external systems.*

- Provides multi-level governance for flexibility

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Information Model and Dependent Artifacts



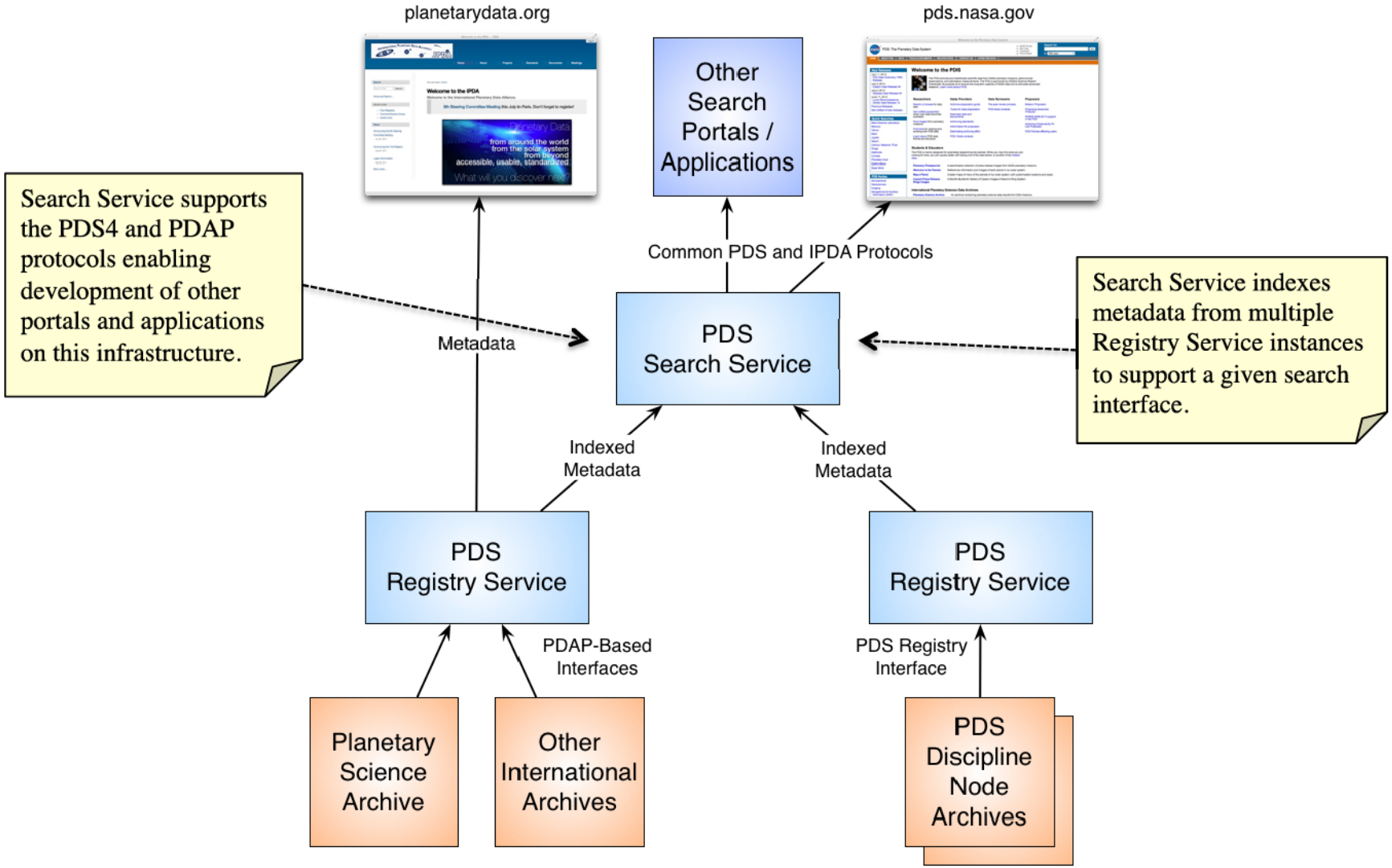
PDS4 Application Architecture

- Service-Based Design
  - *Support remote access to data and services to bring the federation together both for ingestion and distribution.*
- System of Registries
  - *Adopt a system of registries to support improved tracking and access.*
  - *Facilitated by the PDS Registry Service with its REST-based API.*
- Common Search
  - *A publicly available layer facilitating search across PDS.*
  - *Facilitated by the PDS Search Service based on Apache Solr providing support for high performance facet-based search.*
  - *Currently supports the PDS4 Search Protocol and the Planetary Data Access Protocol (PDAP) for data discovery and access.*

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Search Architecture



Virtual Observatory (VO)

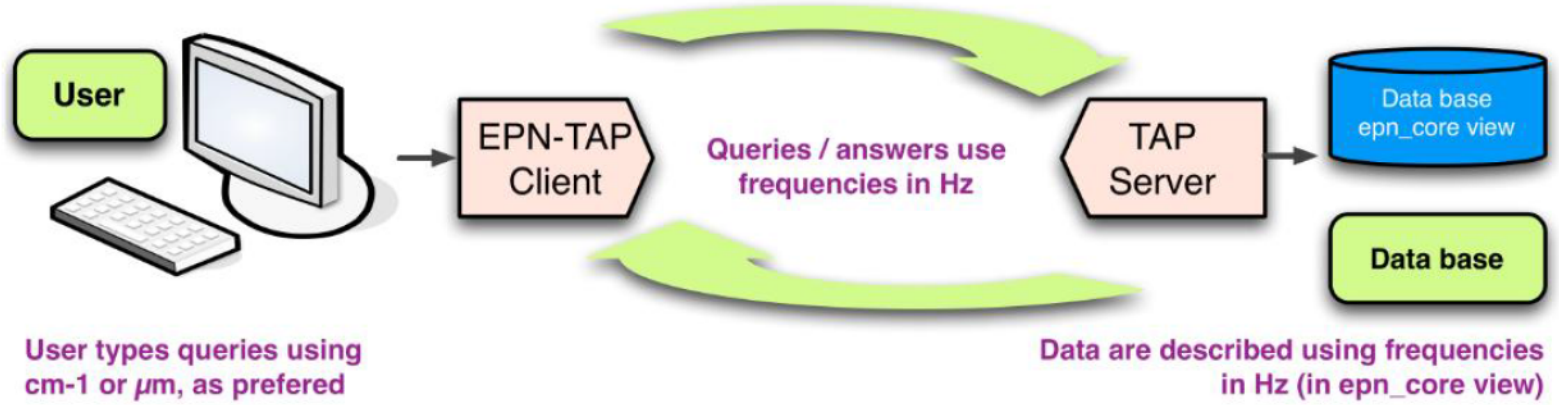
- The Virtual Observatory[1] (VO) is the vision that astronomical datasets and other resources should work as a seamless whole.

- The International Virtual Observatory Alliance[1] (IVOA) is an organization that [*develops consensus]* about the technical standards that are needed to make the VO possible.

[1] http://www.ivoa.net/

EPN-TAP

- The EPN-TAP protocol allows a user to select a subset of data from an archive in a standard way.
  - *Based on the IVOA Table Access Protocol (TAP).*
  - *The TAP mechanism is defined by an underlying data model and reference dictionaries.*



Practical implementation for EPN - TAP queries.[1]

[1]Erard, S.; Cecconi, B.; Le Sidaner, P.; Berthier, J.; Henry, F.; Molinaro, M.; Giardino, M.; Bourrel, N.; André, N.; Gangloff, M.; Jacquey, C.; Topf, F., "The EPN-TAP protocol for the Planetary Science Virtual Observatory", Astronomy and Computing, Volume 7, p. 52-61, 11/2014

Testbed

- A testbed is planned as a part of the continuing International Planetary Data Alliance (IPDA) *Registration and Search* project
  - *Goal is to expose PDS4 registered data and services through the PDS4 search service.*
  - *Search is based on the Apache Solr open source search engine supporting the PDS4 and PDAP protocols.*

- Capability has been demonstrated by the interface to the European Space Agency (ESA) Planetary Science Archive (PSA)
  - *PSA products are searchable through the PDS4 interface.*

Summary

- We have a plan to make planetary science resources available to the international spaces science community
  - *Develop an interface between the PDS and the Virtual Observatory.*
  - *The project is a collaboration of members from the International Virtual Observatory Alliance (IVOA) and the International Planetary Data Alliance (IPDA).*

- A testbed is planned as a part of an IPDA project, IPDA Registration and Search
  - *Integrate the EPN–TAP protocol with the PDS Search Service.*
  - *Plan to report progress at the IPDA 2016 summer meeting.*

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Thank You!

A portion of this research was carried out by the Jet Propulsion Laboratory, managed by the California Institute of Technology under a contract with the National Aeronautics and Space Administration.