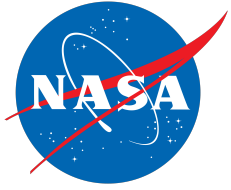




Search, Access and Distribution

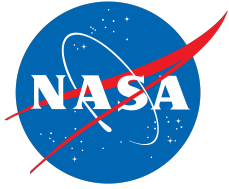
PDS Technical Session
Pasadena, California
September 21-23, 2016

Sean Hardman



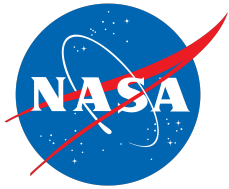
Topics

- Overview
- Lifecycle
- Search Service and its Configuration
- Search Goals
- Service Status/Usage

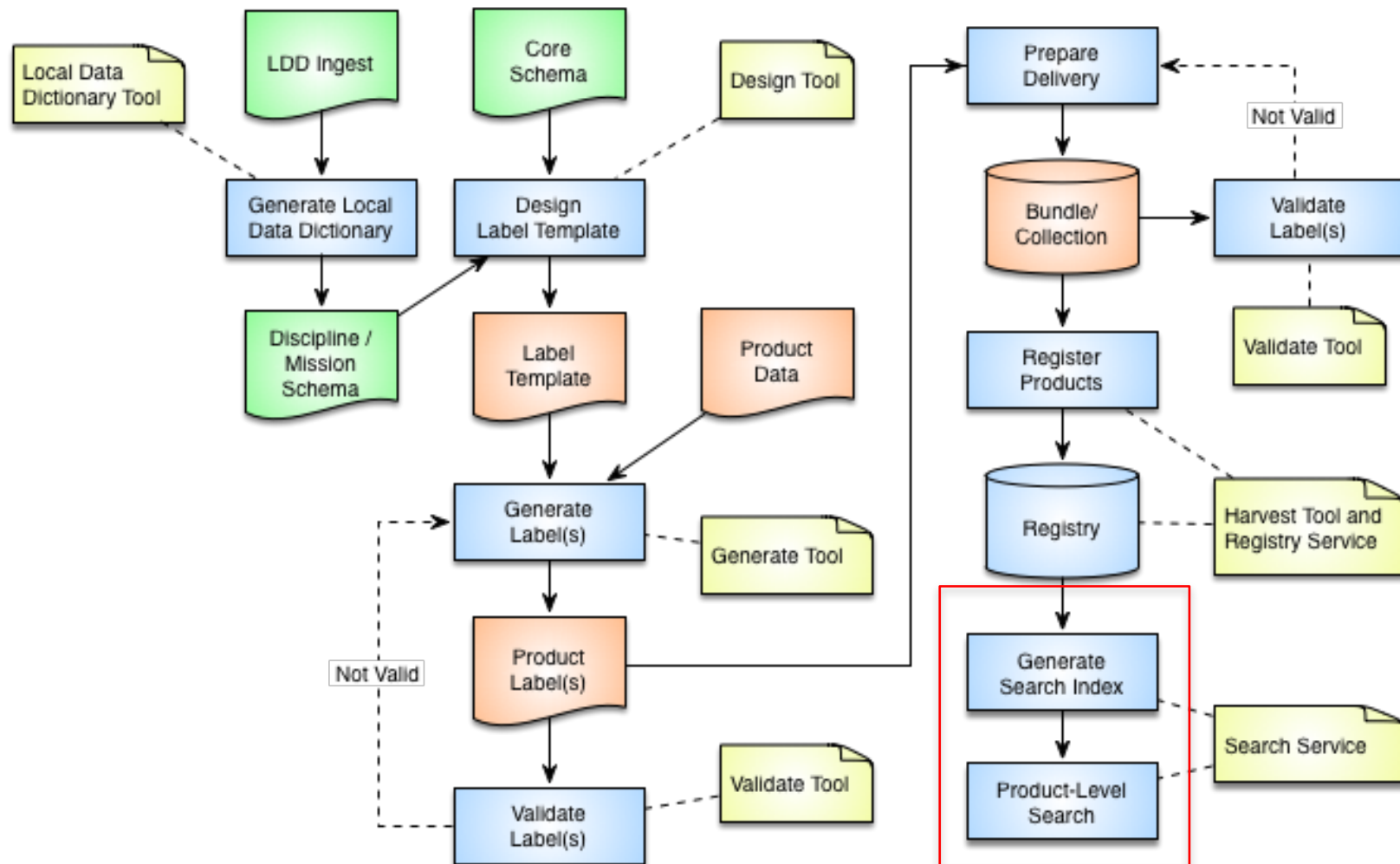


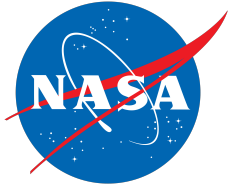
Overview

- Search and access starts with indexing the metadata from the Registry Service.
- The Search Core and Search Service components facilitate index generation and search.
- The Transport components facilitate access to the data products.



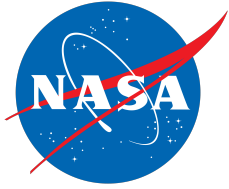
Product Lifecycle





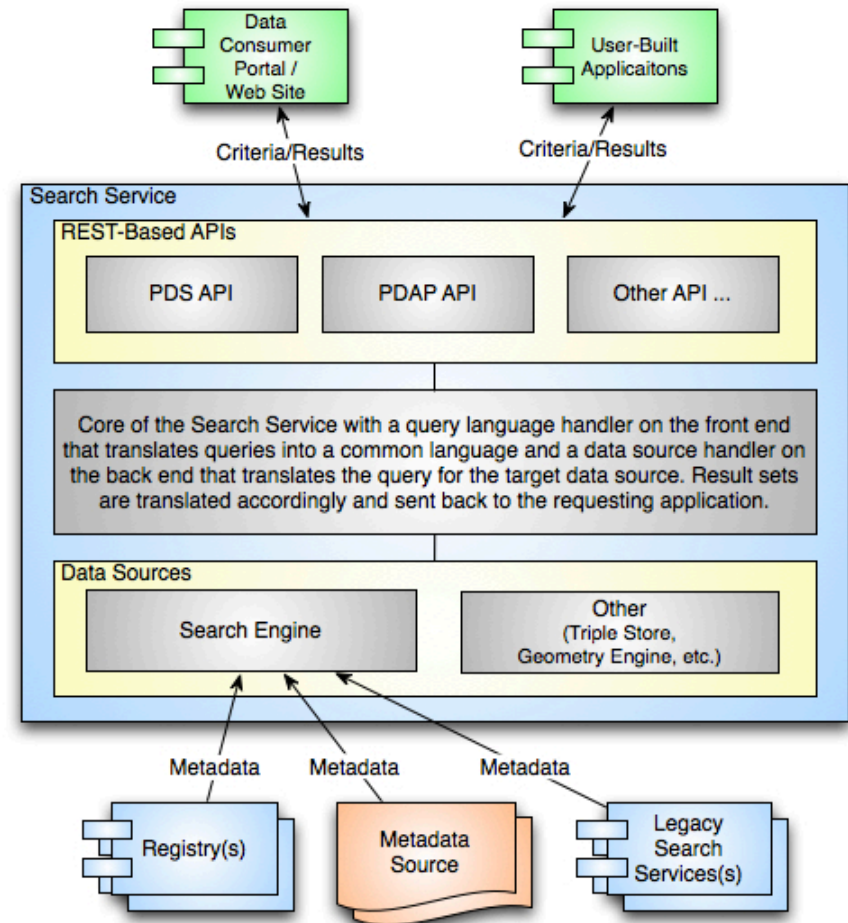
Search Service

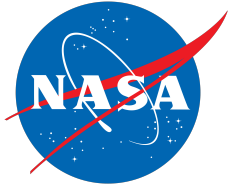
- This service is a deployable component that accepts queries for data and returns a set of matching results.
- Provides the public interface (REST-based over HTTP) to the metadata contained in the federated registries.
- Provides the second line of metadata harvesting within the system in order to facilitate discovery of products.
- Generation of search indices from registry metadata supports multiple query formats and is tailor-able for customized search interfaces.



Service Architecture

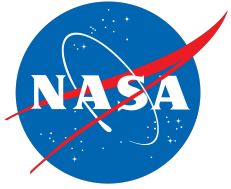
- Search indexes built from multiple sources.
- Allows for annotation of archive metadata.
- Customizable for a discipline-specific search interface.





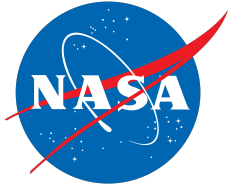
Metadata Annotation

- The data model and the Search Service architecture facilitate metadata annotation.
 - Allows search to be based on the latest and most accurate metadata.
- A defined product with an associated table allows for updated or additional metadata to be specified for a set of products.
 - This product will be associated in the registry with a data set or subset of products.
- The Search Service will support a similar structure for supplying metadata separate from registered content.



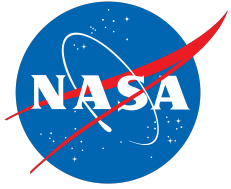
Index Generation

- The main source of metadata for the search index is the contents of the Registry Service.
- The index will make use of the associations in the registry to help users find related products, documents, etc.
- It will also make use of the classification schemes in the registry to help drive the faceted search capabilities.
- Indexes can be tailored for specific search applications.



Search Core Configuration

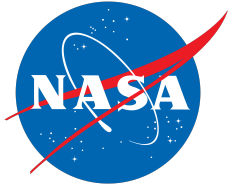
- The tool includes configurations for PDS3, PDS4 and IPDA content.
- Each class of configuration includes a separate configuration file for the product type we intend to index.
- Each class of configuration also includes a properties file specifying the primary and secondary registries to query.



Search Core Configuration Example

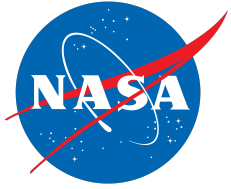


```
<product>
  <specification>
    <title>PDS4-Observational</title>
    <query>
      <registryPath>objectType</registryPath>
      <value>Product_Observational</value>
    </query>
    <query>
      <registryPath>status</registryPath>
      <value>Approved</value>
    </query>
    <checkAssociations>true</checkAssociations>
  </specification>
  <indexFields>
    <!-- Identifier Fields -->
    <field name="search_id" type="required">
      <outputString format="text">pds4:{lid}</outputString>
    </field>
    ...
  </indexFields>
</product>
```



Search Service Configuration

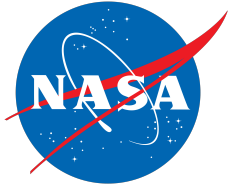
- The service includes the standard Apache Solr configuration tailored to our PDS deployment.
- The PDS-specific portion includes field definitions corresponding to the mapping fields found in the Search Core configuration.



Search Service Configuration Example

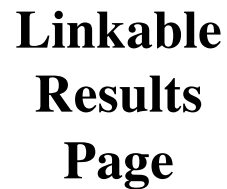


```
<schema name="pds" version="1.4">
...
<field name="search_id" type="string" indexed="true" stored="true" required="true" />
<field name="_version_" type="long" indexed="true" stored="true"/>
<field name="identifier" type="string" indexed="true" stored="true" required="true" />
<field name="abstract_text" type="text_general" indexed="false" stored="true" multiValued="true" />
<field name="agency_name" type="lowercase" indexed="true" stored="true" multiValued="true" />
<field name="alternate_title" type="text_general" indexed="false" stored="true" multiValued="true" />
<field name="archive_status" type="lowercase" indexed="true" stored="true" multiValued="false" />
...
</schema>
```



Search Goals

- Improved integration across nodes and across agencies
 - Pass search parameters to the node search engines (e.g., two/n-tiered search)
 - International data sets showing up in PDS search results
- Better navigation and support for mission and other “virtual” views
 - PDS mission support pages registered in the PDS search infrastructure
 - Facet-based navigation to drive specific views

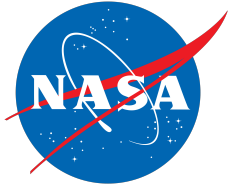


Facet-Based Search

Mission Pages

Data Sets





Mission Pages



Curated to provide user-directed access and support for missions to data, user guides and other information in the PDS.

Quick Searches

- Mercury
- Venus
- Mars
- Jupiter
- Saturn
- Uranus
- Neptune

PDS Web Sites

- PDS
- Atmospheres
- Geosciences
- Imaging
- Navigational & Ancillary Information (NAIF)
- Planetary Plasma Interactions (PPI)
- Planetary Rings
- Small Bodies

PDS Support

- Management
- Engineering

Cassini Archive (Main)

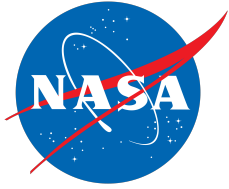
- Jovian Flyby
- Saturnian Bodies
- Fields and Particles

Welcome to the Cassini Archive Page

PRIME	EQUINOX	SOLSTICE
Mar 14, 2004 – July 1, 2008 S01 – S41	July 1, 2008 – Oct. 11, 2010 S42 – S63	Oct. 11, 2010 – Sept. 15, 2017 S64 – S101
100% complete	100% complete	1 st installment delivered October 1, 2011

Cassini data are delivered to the PDS 9-12 months after acquisition and are delivered every three months. PDS must validate those data before they are considered certified. This may result in a delay between the time when Cassini delivers data and the time when PDS makes the data available. Alternately, PDS may choose to put data online while they are being validated, stating that they are not certified. PDS urges caution when using any data that have been released by the PDS for less than three months.

Organization of The Cassini Mission. The mission is divided into 4 phases: the Cruise Phase (including the [Jupiter Flyby](#) Dec 2000-Jan 2001), the Prime Mission (including the [Huygens Probe](#) Landing), the Equinox Mission and the Solstice Mission. Data from all phases are stored sequentially in the PDS. Tables and graphics associated with various aspects will help you scope the mission.



Search for Venus Data

The screenshot shows the PDS Search Results page for the query 'target:venus'. The page is titled 'PDS: The Planetary Data System' and includes a search bar with the query 'target:venus'. The results show 1-50 of 297 results in 0.001 seconds. The left sidebar contains a 'Refine Your Search' section with facets for Model Version (PDS3: 280, PDS4: 11), Agency (ESA: 172, NASA: 118), Type (Data Set: 279, Investigation: 10, Search Tool: 6, Target: 2), Target (Planet: 297, Other: 72, Calibration: 50, Satellite: 16, Comet: 13, Ring: 6, Asteroid: 2), and Investigation (Venus Express: 170). The main content area shows search tools and data sets. The 'Search Tools' section includes links for 'Magellan Image Search' and 'Messenger Image Search'. The 'Data Sets and Information' section lists two data sets: 'VENUS EXPRESS VENUS VRA 1/2/3 NMP 0027 V1.0' and 'VENUS EXPRESS VENUS VRA 1/2/3 EXTENDED MISSION 1 0179 V1.0'.

Refine Your Search

Model Version

- PDS3 (280)
- PDS4 (11)

Agency

- ESA (172)
- NASA (118)

Type

- Data Set (279)
- Investigation (10)
- Search Tool (6)
- Target (2)

Target

- Planet (297)
- Other (72)
- Calibration (50)
- Satellite (16)
- Comet (13)
- Ring (6)
- Asteroid (2)

Investigation

- Venus Express (170)

Search Results

target:venus Search New Search

1-50 of 297 results (0.001 seconds)

Search Tools

These tools let you search for data products matching your query. This is usually the best way to access the data. If no tool looks appropriate, you can browse the matching data sets, below.

Search Tool: Magellan Image Search
Use the Planetary Atlas to search for images from the Magellan mission from Venus or the Earth and Moon encounters.

Search Tool: Messenger Image Search
Use the Planetary Atlas to search for MDIS images from the Messenger mission of Mercury or Venus.

[More...](#)

Data Sets and Information

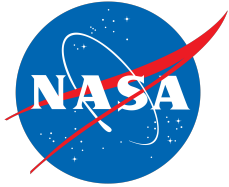
Data Set: VENUS EXPRESS VENUS VRA 1/2/3 NMP 0027 V1.0 (from ESA)
This is a Venus Express Radio Science data set, collected during the Nominal Mission Phase (NMP) 2005-11-09 to 2007-10-02 science subphase 2 (PH2) 2006-07-11 to 2006-09-13. It is a Occultation measurement and covers the time 2006-08-08T01:13:02.050 to 2006-08-08T02:37:40.500.
VENUS EXPRESS - VEX-V-VRA-1/2/3-NMP-0027-V1.0 - starting 2005-12-08T00:08:01.05Z

Data Set: VENUS EXPRESS VENUS VRA 1/2/3 EXTENDED MISSION 1 0179 V1.0 (from ESA)
This is a Venus Express Radio Science data set, collected during the Extended Mission Phase 1 (EXT1) 2007-10-03 to 2009-05-01 science subphase 12 (PH12) 2008-01-04 to 2008-03-31. It is a Occultation measurement and covers the time 2008-02-09T04:45:22.050 to 2008-02-26T05:45:41.950.
VENUS EXPRESS - VEX-V-VRA-1/2/3-EXT1-0179-V1.0 - starting 2007-12-02T00:09:04.05Z

Agency facet allows users to select between ESA and NASA results.

PSA data sets currently link directly to a PSA web-based interface.

Future work includes providing a jump page similar to how PDS data sets are handled.



PSA Interface for Selected Venus Express Data Set



The PSA METADATA Query Ser...

psa.esac.esa.int:8000/aio/jsp/metadata.jsp?DATA_SET_ID=VEX-V-VRA-1

Search

Planetary Science Archive
European Space Agency

The Planetary Science Archive METADATA Query Service

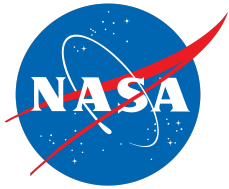
Search Result (1 Data Sets found)

< < Previous Page 1 of 1 Next > >

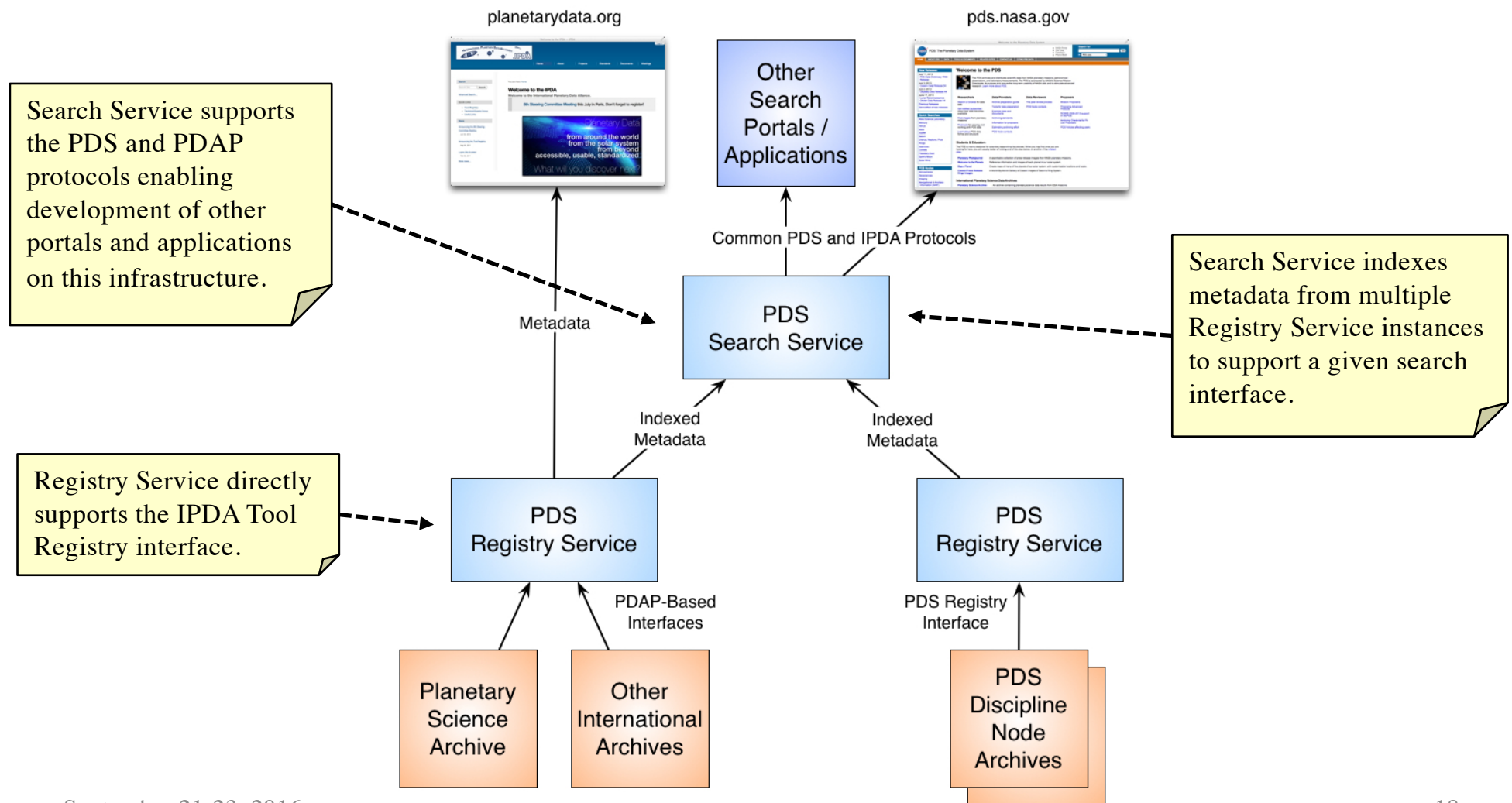
FTP	Files	DATA_SET_ID	DATA_SET_NAME	PRODUCTS	INSTRUMENT_ID	TARGET_NAME	START_TIME	STOP_TIME
FTP	XML/Tree	VEX-V-VRA-1/2/3-NMP-0027-V1.0	VENUS EXPRESS VENUS VRA 1/2/3 NMP 0027 V1.0	232	VRA	VENUS	2006-08-08 01:13:02.05	2006-08-08 02:37:40.5

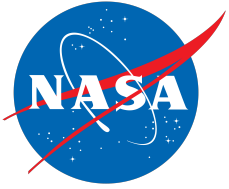
< < Previous Next > >

Total 1 DataSets



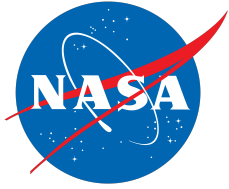
Search Architecture





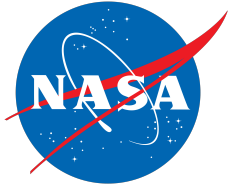
Service Status

- The Search Service instance deployed for PDS.
 - Search web-based interface accessible at:
<http://pds.nasa.gov/tools/data-search/>
 - Search REST-based API accessible at:
PDS Protocol:
<http://pds.nasa.gov/services/search/search/>
PDAP Protocol:
<http://pds.nasa.gov/services/search/pdap/>



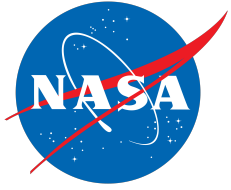
Service Usage

- The Search Service can be queried for the purpose of content discovery and incorporation into other web interfaces.
 - PDS instance currently contains the PDS3/PDS4 catalog objects, PDS3 data sets, PDS4 bundles/collections and PSA data sets.
- Local deployment of the above component is encouraged to support local search scenarios.



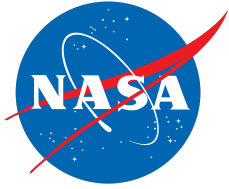
Access and Distribution

- Now that we can discover data from the Search Service, how do we access it.
- To facilitate access and distribution, we developed the Transport Services.
- Yes, there are two different implementations.
- We may merge them at some point in the future but they are pretty different in their approaches.
- Both services offer REST-based interfaces.



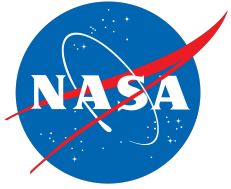
Transport Service (OFSN)

- OFSN stands for Online File Specification Name.
- This implementation of the service is very similar in functionality to the PDS-D Product Server.
- Along with support for all of the old PDS3 return types (e.g., DIRLIST, etc.), it also supports several of the transformations offered by the Transform Tool.



Transport Service (Registry)

- Although it has “Registry” in the name, the service supports querying both the Registry Service and the Search Service for finding products.
- The service allows one or more logical identifiers to be passed in and the service returns all files associated with each product in a ZIP or TAR/GZIP package.
- This service does not yet offer transformations.



Service Status/Usage

- The Transport Service instances deployed for PDS.
 - <http://pds.nasa.gov/services/transport-registry/>
 - <http://starbase.jpl.nasa.gov/services/transport-ofsn/>
- Both services return PDS3/PDS4 context products.

Questions/Comments