



PDS Portal and Search

PDS Virtual Tool Summit
February 23, 2015

Sean Hardman and Emily Law



Topics



- Approach
- Architecture
- Search Service
- Search Protocol
- EN Search Interfaces
- Wrap Up



Approach

- Build the software system based on generic common software and common protocols for accessing that software.
 - PDS Registry Service with its REST-based API is the main component.
 - PDS Search Service based on Apache Solr provides support for high performance facet-based search.
- Utilize the PDS4 data model for data object definitions and to configure the software where appropriate.
 - The model defines the key context objects (i.e., Data Set, Instrument, etc.).



Query Models

- Subsets of the information model
 - Identifies attributes whose values
 - to be harvested from labels or data files
 - used as search parameters for
 - Text- or Field-based search
 - Facet-based search
 - E.g. Product_Observational
 - Mission Science Data Collections
- Specified in version of a Local Data Dictionary

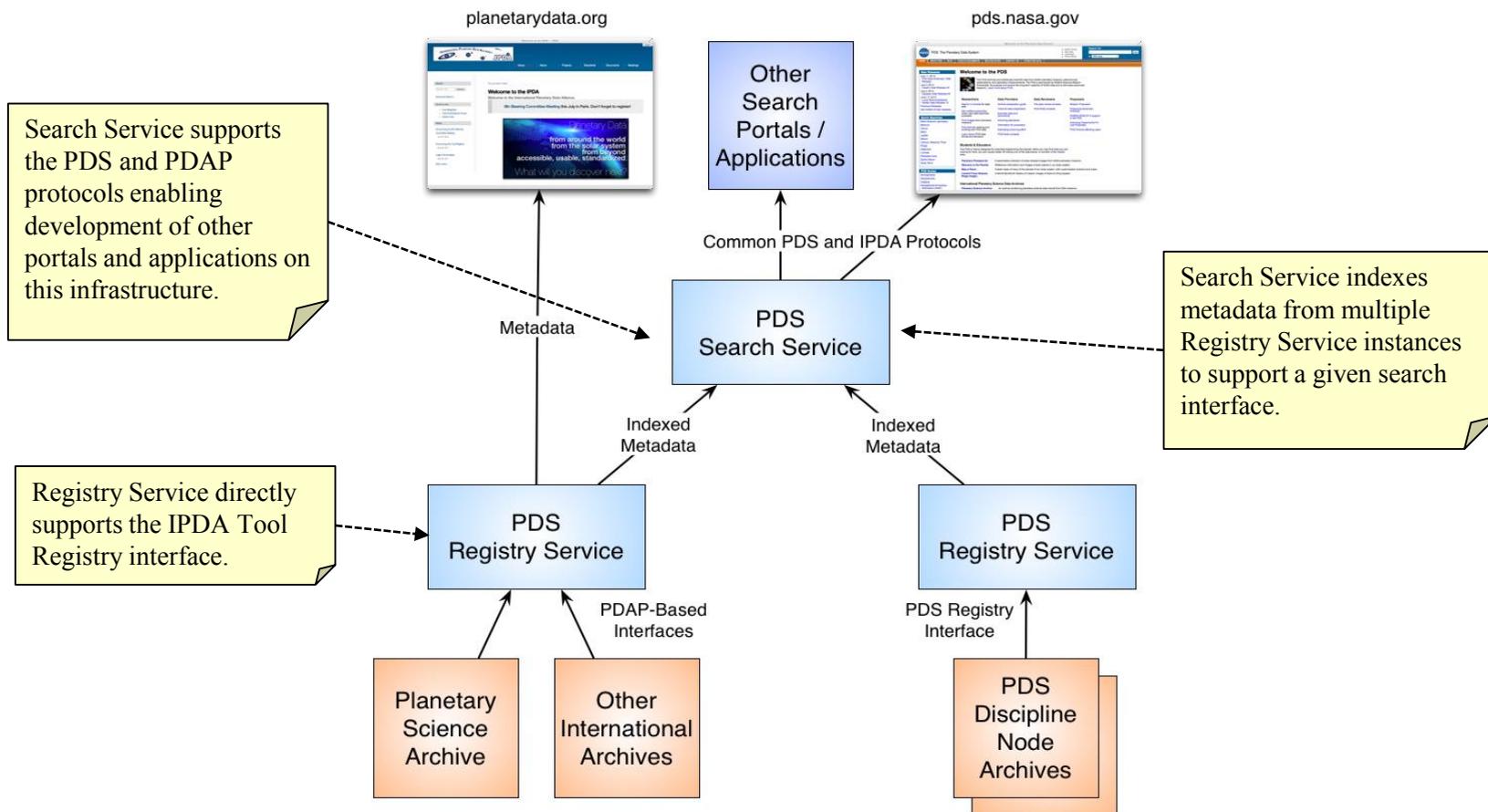


Using the Models

- Core Search should be fully driven by the information model
- DN Search should take advantage of the information model
- EN/DN need to identify a common formats for query models
- If there are gaps or other useful information needed, we can add them



Architecture





Search Service

- This service is a deployable component that accepts queries for data and returns a set of matching results.
- Using Apache's Solr for the search engine portion of the service.
- 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.



Search Protocol

- PDS search protocol implemented as a REST-based interface over HTTP.
 - <http://pds.nasa.gov/services/search/search?target=mars>
- Supports return of paged results in a defined structure (e.g., XML or JSON).
- Implementation of this protocol across PDS facilitates parameter passing and integration.
- The architecture allows support for other protocols.
 - For example, IPDA's Planetary Data Access Protocol (PDAP)



Search Via REST-Based Interface



- Navigate to the API endpoint:
 - [http://pds.nasa.gov/services/search/search?term="venus express"&return-type=xml](http://pds.nasa.gov/services/search/search?term='venus express'&return-type=xml)
- The following should appear in the browser:

```
<response>
  <lst name="responseHeader">
    <int name="status">0</int>
    <int name="QTime">7</int>
    <lst name="params">
      <str name="term">"venus express"</str>
      <str name="return-type">xml</str>
    </lst>
  </lst>
  <result name="response" numFound="533" start="0" maxScore="1.47908">
    <doc>
      <float name="score">1.47908</float>
      <str name="title">VENUS-EXPRESS VENUS MAG 4 EXTENSION2 V1.0</str>
    ...
```



Catalog-Level Search

The screenshot shows the PDS Search Results page for the query 'clementine'. The page includes a search bar with the query 'clementine' and a 'Go' button. Below the search bar, there are navigation tabs for 'HOME', 'ABOUT PDS', 'DATA', 'TOOLS & DOCUMENTS', 'RELATED SITES', 'CONTACT US', and 'CITING PDS DATA'. The 'DATA' tab is selected. The main content area is titled 'Search Results' and shows '1-35 of 35 results (0.006 seconds)'. There are three main sections: 'Search Tools', 'Data Sets and Information', and 'Collection: Clementine 750nm UVVIS Basemap Digital Image Model Basemap Mosaic - Data'. The 'Search Tools' section lists 'Search Tool: Clementine Image Search' and 'Search Tool: Lunar Orbital Data Explorer'. The 'Data Sets and Information' section lists 'Collection: Clementine 750nm UVVIS Basemap Digital Image Model Basemap Mosaic - Data', 'Bundle: Clementine 750nm UVVIS Basemap Digital Image Model Basemap Mosaic', and 'Collection: Clementine 750nm UVVIS Basemap Digital Image Model Basemap Mosaic - XML Schema'. The 'Collection: Clementine 750nm UVVIS Basemap Digital Image Model Basemap Mosaic - Data' section provides details about the data collection, including the wavelength (750-nm) and the camera used (Ultraviolet/Visible Camera).

Search facets tailored to the catalog-level metadata.

Direct links to search tools matching the query criteria.

Search results including PDS3 and PDS4 data.



Product-Level Search

The screenshot shows the PDS Search Results page. At the top, there is a search bar with the text 'Search for:' and a 'Go' button. Below the search bar, there is a navigation menu with links for HOME, ABOUT PDS, DATA, TOOLS & DOCUMENTS, RELATED SITES, CONTACT US, and CITING PDS DATA. The 'DATA' link is highlighted. Below the navigation menu, there is a search bar with the text 'Search for:' and a 'Go' button. Below the search bar, there is a dropdown menu with the text 'in PDS data'. The main content area is divided into two sections: 'Refine Your Search' and 'Search Results'. The 'Refine Your Search' section has four categories: Type, Investigation, Instrument, and Processing Level. The 'Search Results' section shows the search term 'met' and the number of results '1-22 of 22 results (0.007 seconds)'. The results are grouped into 'Bundles and Collections' and 'Products'. The 'Bundles and Collections' section lists several bundles and collections, including 'MET Bundle', 'LIDAR Bundle', and 'SSI Bundle'. The 'Products' section lists several observational products, including 'PHOENIX MARS MET Experiment'.

Search facets tailored to observational product metadata.

Results list tailored as well, focusing on Bundles, Collections and Products.



Wrap Up

- The Registry and Search Services continue to be tuned.
- Continue to add support for additional PDS4 products in the Search Service.
 - Recently added support for dictionary related products.
- Future work includes expanding the search protocol(s).

Questions/Comments