Title
Enabling interoperability in planetary sciences and heliophysics: The case for an information model

Journal
PLANETARY AND SPACE SCIENCE, 150

ISSN
0032-0633

Authors
Hughes, J Steven
Crichton, Daniel J
Raugh, Anne C
et al.

Publication Date
2018

DOI
10.1016/j.pss.2017.04.005

Peer reviewed

Planetary and Space Science 150 (2018) 43–49

Contents lists available at ScienceDirect

Planetary and Space Science

journal homepage: www.elsevier.com/locate/pss

Enabling interoperability in planetary sciences and heliophysics: The case for an information model

J. Steven Hughes [a,*], Daniel J. Crichton [a], Anne C. Raugh [b], Baptiste Cecconi [c], Edward A. Guinness [d], Christopher E. Isbell [e], Joseph N. Mafi [f], Mitchell K. Gordon [g], Sean H. Hardman [a], Ronald S. Joyner [a]

[a] *Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA*
[b] *University of Maryland, College Park, MD, USA*
[c] *Baptiste Cecconi, Paris Observatory, Paris, France*
[d] *Washington University, St Louis, MO, USA*
[e] *USGS Astrogeology Science Center, Flagstaff, AZ, USA*
[f] *Institute of Geophysics and Planetary Physics, UCLA, Los Angeles, CA, USA*
[g] *SETI Institute, Mountain View, CA, USA*

ARTICLE INFO

Keywords:
Interoperability
Information architecture
Information model
Planetary science
Heliophysics
Semantic
Digital repository
Ontology

ABSTRACT

The Planetary Data System has developed the PDS4 Information Model to enable interoperability across diverse science disciplines. The Information Model is based on an integration of International Organization for Standardization (ISO) level standards for trusted digital archives, information model development, and metadata registries. Where controlled vocabularies provides a basic level of interoperability by providing a common set of terms for communication between both machines and humans the Information Model improves interoperability by means of an ontology that provides semantic information or additional related context for the terms. The information model was defined by team of computer scientists and science experts from each of the diverse disciplines in the Planetary Science community, including Atmospheres, Geosciences, Cartography and Imaging Sciences, Navigational and Ancillary Information, Planetary Plasma Interactions, Ring-Moon Systems, and Small Bodies. The model was designed to be extensible beyond the Planetary Science community, for example there are overlaps between certain PDS disciplines and the Heliophysics and Astrophysics disciplines. "Interoperability" can apply to many aspects of both the developer and the end-user experience, for example agency-to-agency, semantic level, and application level interoperability. We define these types of interoperability and focus on semantic level interoperability, the type of interoperability most directly enabled by an information model.

1. Introduction

The Planetary Data System (PDS) is NASA's planetary science data archive and has the mission to provide the near-term discoverability, long-term preservation, and usability of the data returned from all NASA supported missions to explore the solar system. The digital repository currently contains about two petabytes of data from across a diverse set of science disciplines. The data was collected from over 1200 instruments, including both remote and in situ measurements.

After its initial release in 1990 and two decades of operations the PDS initiated development of its next generation system, PDS4. Based on lessons-learned, it is a complete redesign of the PDS to develop a system that meets the demands of higher data volume and to leverage new information technologies. The PDS4 architecture has two primary components: the information architecture and the software/technical architecture. The PDS4 Information Architecture (Hughes et al., 2009, 2014; Crichton et al., 2011) defines the system's informational requirements and institutes a multi-governance scheme for management of the architecture components. The PDS4 Information Model (Hughes et al., 2009, 2014), the core component of the architecture, is based on the ISO 14721 (ISO 14721, 2003) and ISO/IEC 11179 standards (ISO/IEC 11179, 2008). It was designed to be extensible across science disciplines and to promote interoperability that facilitates data sharing internationally.

The PDS4 Information Model provides a hierarchical structure for data archiving with three types of products. The Bundle Product is a list of related collections. The Collection Product is a list of related basic

* Corresponding author.

https://doi.org/10.1016/j.pss.2017.04.005
Received 15 December 2016; Received in revised form 25 February 2017; Accepted 6 April 2017
Available online 12 April 2017

products of similar type, for example the spectral cubes from a single instrument. The Basic Product is the smallest unit of data registered and tracked under PDS4. A Basic Product may consist of an image but may also include supplemental information such as a related engineering table. The model defines four fundamental data structures: Array - a homogeneous n-dimensional array of scalars (e.g., images or spectral cubes); Table – the traditional fixed-width structure based on a single record with heterogeneous binary or character fields; Parsable Byte Stream – a stream from which the data value can be extracted directly by applying parsing rules to the bytes (e.g., simple text files, XML files, CSV tables); Encoded Byte Stream – a stream in which the bytes must interpreted, transformed, or otherwise processed before the data can be extracted (e.g., PDF files, JPEG images, MPEG movies) (Hughes et al., 2015; Raugh and Hughes, 2015).

The software/technical Architecture is a distributed service-oriented architecture encompassing the individual PDS discipline nodes and the PDS's international partners. The architecture provides consistent protocols for access to the data and services and a federated registry infrastructure to track and manage the contents of the digital repository. The current distributed search infrastructure is based on metadata harvested from product labels and loaded into Apache Solr (Apache Solr), an open source enterprise search platform.

PDS4 is the first operational science information system resulting from an information model-driven development methodology (Crichton et al., 2014). It is being used to coordinate data archiving in both the national and international planetary science communities. With the system's information requirements captured in an ontology modeling tool significant but controlled change can occur as the science domains and implementation technologies change.

The PDS4 Information Model enables interoperability across the Planetary Science and related space science disciplines. In general the term interoperable is or relates to the ability to share data between different computer systems. In the following more specific aspects of interoperability are described.

1.1. Agency-to-Agency level

At the Agency-to-Agency level independent systems do not share a common infrastructure but are interoperable because of a mutual interest in the information products.

This type of interoperability is supported by the underlying standards. A good example is the interface between PDS and its deep archive, the NASA Space Science Data Coordinated Archive (NSSDCA). The "information package", the information stored by the archive (ISO 14721, 2003), provides the interoperability link needed to connect the two and support this vital relationship. Commonality of structure and metadata concepts shared by both institutions simplifies the transfer of information and the core operations of the target (NSSDCA) process.

Another good example of agency-to-agency level interoperability is the interface between the PDS and the European Space Agency (ESA) Planetary Science Archive (PSA) (Besse et al., 2016). This interface enables queries between the two systems using the Planetary Data Access Protocol (PDAP) (IPDA Planetary Access Protocol, 2013) and a common set of keywords to ensure mirroring of resources.

Interfacing with other archives built on the same standards is accommodated by the common terminology and structural skeleton defined by the standards.

1.2. Semantic level

At the semantic level systems interoperate based on the commonality of definitions of key concepts. These common definitions present an interface between the systems. The common definitions can also be viewed as shared knowledge.

The development of the PDS4 Information Model (IM) and its partition into discipline namespaces is an application of this. The model-driven design paradigm prevents unintentional bifurcation of meaning and supports partitioning of the model into namespaces that can be mapped directly to and managed as distinct contexts. A namespace provides a unified set of attributes to define something like display orientation in all product contexts in which the concept is applicable. The common definition provides the basis for programmatic interoperability by providing developers with a single reference point for display information. And that, in turn, enables applications and other namespaces to take advantage of the established terminology to, for example, describe target orientation within a displayed image.

1.3. Application level

At the application level, the systems support interactions between disparate systems and make the interactions look seamless from the end user's perspective.

The EuroPlaNet (EPN) Table Access Protocol (TAP) (Erard et al., 2014) interface and Virtual European Solar and Planetary Access (VESPA) (Erard et al., 2015) projects are good examples of this - adding a software layer between application and target archive that allows a user to treat products from disparate sources as computationally equivalent. The PDS4 service structure and its Application Program Interfaces (APIs) are designed to support this sort of interoperability, and the PDS4 Information Model can support the semantic translation mapping needed to interface the PDS4 named concepts to those in the target environment.

2. A brief history of interoperability in the space sciences

Since before the advent of the World Wide Web, interoperability across space science digital repositories has been a goal. In 1982 and 1986 the Committee on Data Management and Computing (CODMAC) issued reports that set guidelines for the development of science data archives (Bernstein et al., 1982; Arvidson et al., 1986). The committee recommended that sufficient ancillary and metadata be captured and archived with the data to ensure that future users of the data would be able to understand how to interpret the science data formats as well as understand the context under which the data was collected and processed.

The Planetary Data System (PDS) was established in 1989 based on CODMAC principles. In 1999, after the advent of the World Wide Web, PDS deployed the PDS Distributed Inventory System (DIS) (Hughes and McMahon) which harvest metadata from PDS product labels and provided a product location and retrieval services across the PDS's heterogeneous and distributed nodes. Also in 1999, the Interoperable Systems for Archival Information Access (ISAIA) team (Hanisch, 1999) was formed as a collaboration of several space science repositories, including the PDS, with the ambitious goal to provide an "interdisciplinary data location and integration service for space science" (Hanisch, 2002). The importance of metadata standards was highlighted.

In 2001, Uschold (Uschold and Gruninger, 2004) argued that a "single shared ontology" is critical for developing a digital library that enables semantic interoperability across disciplines. And in 2002, the report prepared by the National Virtual Observatory Science Definition Team (Hanisch, 2002) emphasized standards for metadata and data formats for accessing large astronomical data sets.

The Space Physics Archive Search and Extract (SPASE) (Thieman et al., 2010), an international consortium formed in 2001, is a community-based effort to define standards and services for the Space and Solar Physics community focusing on the Heliophysics data environment. The goals were to define a data model for Space Physics and enable interoperability in a distributed environment to allow resources to be easily registered, found, accessed, and used. The SPASE Metadata Model is an information model for describing the elements of the Heliophysics data environment.

Over the ensuing decade and a half, there have been many successful efforts where shared knowledge has been collected in support of data access and interoperability. The PDS is represents a successful effort for a large and diverse interdisciplinary community. There have been many lessons learned (Hughes et al., 2015) for example:

- There is never a definitive, exhaustive source; and it is not uncommon to find contextual nuance at work in the use of discipline terms that are considered ‘well-known’.
- The shared knowledge is almost impossible to manage as a single monolithic unit because of the disparate sources but should be partitioned in order that a multi-level governance scheme can be applied.
- The knowledge to be shared must be collected in a formal language using accepted standardized modeling methodologies otherwise inconsistencies and ambiguities will over time significantly degrade the effectiveness of the knowledge.
- The stability of an information system is highly dependent on the stability of the shared knowledge. But at the same time, the shared knowledge must evolve to remain relevant as the science discipline evolves over time.

3. Overview of the PDS4 information model

The PDS4 Information Model[1] captures the knowledge about the planetary science community’s digital repositories at several levels of specificity and provides a means by which both humans and machines can “communicate” about the digital content of the repositories. As a set of formal definitions, the model is also leveraged as a set of information requirements for the PDS4 Information Architecture (Crichton et al., 2014). Multi-level governance, depicted in Fig. 1, is instituted by the model at the common, discipline, and mission levels to enable interoperability at the appropriate level and simplify the management of the model as it evolves over time.

At the common or upper level of an information model resides the knowledge about what “things” (digital objects and products, in the case of the PDS archive) can be located and retrieved and how they are identified, referenced, and packaged. Digital objects in the repository must also have representation information provided in logical and well-defined terms so that they can be properly interpreted for scientific studies. At the next level, shared knowledge in specific disciplines must be available to understand and advance science, for example standard geometry models are needed to determine location and standard cartography models are needed for maps. Finally, a standard vocabulary is required within individual teams to communicate and effectively support the investigation.

Since its release, the PDS4 Information Model and the resulting PDS4 Data Standards have been successfully used by three planetary science missions to archive their data, Lunar Atmosphere and Dust Environment Explorer (LADEE), Mars Atmosphere and Volatile EvolutioN (MAVEN), and Balloon Observation Platform for Planetary Science (BOPPS).

The PDS4 Information Model and the resulting PDS4 Data Standards have also been recommended as the de-facto planetary science data standard by the International Planetary Data Alliance (IPDA). The European Space Agency (ESA) Planetary Science Archive (PSA), a member of the IPDA, supports “the two formats of the Planetary Data System (i.e., PDS3 and PDS4), as well as providing new ways for searching the data products with specific metadata and geometrical parameters.” (Besse et al., 2016) The PSA already uses PDS3 for several missions (SMART-1, Huygens, Mars Express, Venus Express and Rosetta), and PDS4 for the current ExoMars mission; future missions, including BepiColombo, will use PDS4 as well. The IPDA has an ongoing PDS4 implementation task staffed by PDS and PSA personnel and PDS4 training sessions have been held.

[1] An information model in data engineering is a representation of concepts and the relationships, constraints, rules, and operations to specify data semantics for a chosen domain of discourse. It provides a sharable, stable, and organized structure of information requirements or knowledge for the domain context. Lee (1999)

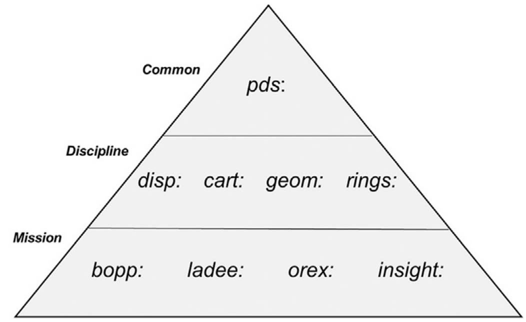

Fig. 1. - Multi-level Governance in the PDS4 Information Model.

4. Foundational principles

To provide a stable foundation for the PDS4 Information Model and to keep it as broadly applicable as possible, several established meta-models were adopted as illustrated in Fig. 2. The ebXML federated registry model (Organization for the Advancement of Structured Information Standards, 2005) provided essential concepts including the registry object and object classification, identification, and versioning. Defining these concepts in the model allows a registry to be configured by the model for a specific science discipline. The Open Archival Information System (OAIS) Reference Model (ISO 14721, 2003) provided the core concepts of the digital object, information object, information package, and a metadata classification scheme. The ISO/IEC 11179 (ISO/IEC 11179, 2008) metadata registry reference model provided a comprehensive data dictionary schema that is necessary for defining science terms. For example the values in an enumerated list must be accompanied by a value meaning and terminological entries are allowed for names and definitions in other natural languages.

In the PDS4 Information Model each “thing of interest” in the community, for example a planetary image, is defined to the extent necessary to meet the functional requirements of the system. These definitions comprise the *Domain Knowledge* depicted in Fig. 2. For example, sufficient representation information must be available for a digital object such as the planetary image to be interpreted and used in science research. In comparison, the information about the mission that collected the digital object may be limited to its name, a short description, and references to supporting documentation.

A fundamental principle used in the development of the PDS4 Information Model is that the model remains independent from its implementation. In a classic enterprise architecture such as that presented in the Zackman Framework of Enterprise Architecture (Zachman, 1987) the architecture is partitioned into architectural elements, for example “why”, “how”, “what”, “who”, “where” and “when”. Each element is then modeled at five levels, from contextual to the detailed. The PDS4 Information Model encompasses the “what” element of the architecture, that is, the data being processed or archived. The model is agnostic to the other elements, especially the “how”, or the implementation element of the architecture. A model that is independent of the implementation is inherently more stable because it can more readily change as the science discipline changes. Concurrently it is shielded from technology changes which typically changes at more rapid pace.

To further manage complexities during the developmental and evolutionary phases of the PDS4 Information Model, the multi-level governance scheme is instituted at the common, discipline and mission levels. The common model is governed under a formal change control process where a change control board (CCB) decides whether to approve each change request based on the change’s potential impact on the overall PDS enterprise. At the discipline and mission levels similar governance but with contextually limited scope are instituted.

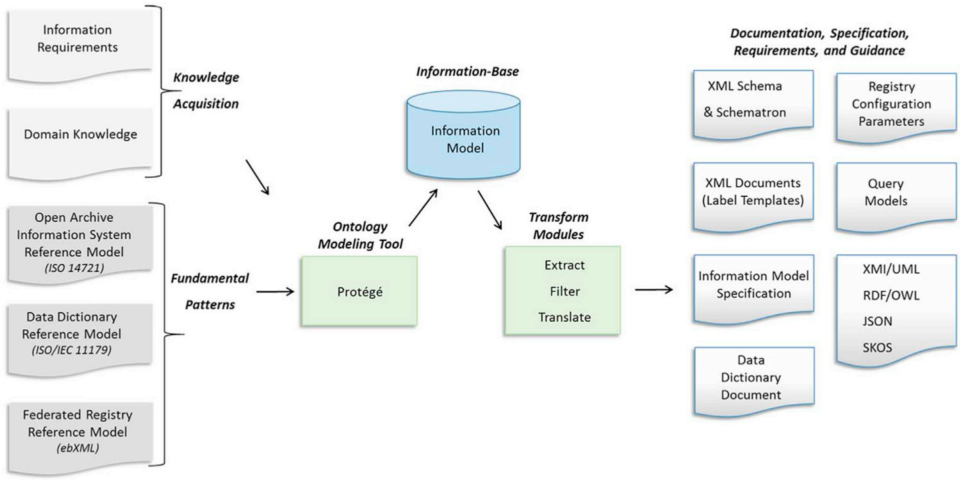

Fig. 2. - PDS4 Information Model.

For implementation, the contents of the PDS4 Information Model are extracted and translated to XML Schema (XML Schema Part 2 1 2 1 1, 2004; XML Schema Part 2 1 2 1 2, 2004) and Schematron files (ISO/IEC Information Technology, 19757-3) as illustrated in Fig. 2. The XML Schema files are used to create XML (Extensible Markup Language (XML) (2008)) documents. Data providers subsequently populate the documents and validate the results using the XML Schema and Schematron files. The XML files are used to label the digital objects in the repository.

Version 1.0 of the PDS4 Information Model was released in 2012 (Hughes et al., 2012). A six month build cycle was established to provide a predictable and stable development schedule as the model continued to mature and discipline development began in earnest.

5. Discipline level models

To remain relevant in an ever changing science discipline; and in support of interoperability, the design process expected discipline extension to happen immediately, and the same development principles and tools are applied to the discipline-level and mission-level development as to the common part of the model. Having overarching discipline dictionaries to address themes like geometric metadata or display orientation metadata is a key way for PDS to ensure interoperability of tools on its own holdings. Several discipline level models have been or are currently under development:

5.1. Display model

The PDS4 display dictionary is a cross-discipline dictionary that is designed for use with image data that apply to many planetary science disciplines. The role of the PDS4 display dictionary is to provide a common method to map two spatial dimensions of an image array to the vertical and horizontal directions of a display device. The dictionary also contains attributes to specify which axis of a 3D image array should be used to as color bands such that a set of planes within that axis can be mapped to blue, green, and red channels of a color display.

5.2. Geometry model

The PDS4 geometry discipline model was developed to capture observational geometry metadata for planetary data. The geometry model provides a set of parameters that describe the conditions with which raw observational data are acquired and are thus important ancillary information for planetary data sets archived by the PDS. Accurate and well-defined geometry metadata is essential for processing the raw observational data into calibrated and derived data products, e.g., calibrated and map projected products. Geometry information captured in the PDS4 discipline model includes, for example, attributes for lighting and viewing angles, for position and velocity vectors of a spacecraft relative to the Sun and relative to the observing body at the time of an observation, and for the location and orientation of an observation projected onto the surface of a target body.

The content of the PDS4 geometry model has been developed from requirements gathered from domain experts in the planetary science community including researchers and data producers. The PDS4 geometry model provides consistency in geometry metadata incorporated in PDS4 data products across the wide range of planetary science disciplines and data collected by instruments observing many types of solar system bodies such as planets, ring systems, moons, comets, and asteroids. It also standardizes the definitions for the geometry attributes across PDS4 archives. Where there is overlap in concepts and terminology, the geometry model is consistent with usage in the PDS4 cartography model.

PDS also provides geometry data in the form of SPICE kernels that are archived by the PDS Navigation and Ancillary Information Facility (NAIF) node. SPICE kernels contain geometry data in the form of ephemerides for spacecraft and solar system bodies, along with instrument parameters and pointing as a function of time. In other words, SPICE data does not refer to particular observations, but can be used to derive observation geometry for such observations. The geometry information in a data product's PDS4 label is the specific geometry at the time or time range that the observation was made.

5.3. Cartography model

Generation and use of cartographic products are essential in support of scientific exploration and research. The PDS Cartography and Imaging Sciences node has lead a coordinated effort toward development of a discipline level Cartographic model in compliance with the primary PDS4 Information Model. The initial cartographic implementation utilized an existing terrestrial Federal Geographic Data Committee (FGDC) geospatial standard (Content Standard for Digital Geospatial Metadata, 1998). For PDS4, the FGDC standard has been extended and adapted in satisfying planetary requirements. For example, extension of existing and creation of new attributes and elements are required in order to describe tri-axial (implemented) and irregular shaped bodies (in progress, in coordination with geometry efforts), define map projection coordinate offset and origin parameters, and specify other unique parameters specific to planetary mapping needs. Implementation of these standards enables the scientific community the means of describing cartographic products used within planetary mapping and research, and satisfies short and long term usability and preservation requirements. In addition, PDS standards development, cartography or otherwise, intends to provide product and metadata information that can be utilized once ingested within user applications beyond the PDS (e.g. Geographic Information System (GIS) or other mapping and database applications). Current and future utilization of the cartographic model across Mission and PDS archive activities will likely reveal additional requirements, and by design, influence intentional evolution of the model.

5.4. Planetary plasma interactions models

The PDS/PPI node has developed a number of separate discipline models in order to enhance the information that PDS metadata are able to provide for CDF-formatted data files, which are the primary archival format for many of the MAVEN instruments. The data in a CDF file is stored in the form of single- or multi-dimensional arrays. While it is possible to describe these structures using the PDS core dictionary, there were no means of describing the logical relationships which exist between the various data arrays. In order to provide this information the PPI node created the following discipline level discipline models:

- Particle discipline model – This dictionary defines a series of the three classes which describe the relationship between the data array, and other supplemental arrays within a CDF file. Axis_Values identifies a 1-D array containing data associated with a single axis of the data array. Face_Values identifies a multi-dimensional array containing data associated with a "face" (i.e. multiple axes) of the data array. Aligned_Values identifies an array with the same dimensionality and axes as the data array which contains data which are supplemental to the data array values (e.g. uncertainties, etc.).
- Alt discipline model – This dictionary defines the Alternate_Values class which identifies arrays with the dimensionality that are equivalent in function and may be used interchangeably. An example would be multiple time columns in a single data file.

CDF allows for scalar values to be stored as data in form of a 1-D array with a single value. These values are captured in the PDS metadata using the "Parameters" class, which is defined in the MAVEN Mission discipline models.

The PDS/PPI node is concurrently a member of the Heliophysics community and has had an integral role in the development of the PDS4 Information Model. They have developed Heliophysics specific models and definitions for digital objects from the MAVEN mission.

5.5. Ring-moon systems models

The PDS4 rings discipline model was developed to capture ring specific supplemental metadata, and observational geometric metadata additional to that provided by the geometry discipline model. Fully describing ring observations is complicated. Each ring is composed of individual particles whose orbital velocities vary based on their radial separation from the central body. A ring may be inclined with respect to the equator plane of the central body and/or with respect to the other rings within the system. The local surface density, optical depth, and light scattering properties within a ring vary due to interactions at specific radial locations with satellites in the system, and ephemeral sub-kilometer structures (wakes) within the rings which are the result of gravitational interactions between ring particles.

The rings local dictionary currently contains more than 70 attributes, subsets of which support general ring system observations and observations of occultations of various sources by ring systems. There are three types of ring occultation observations, stellar – when the ring system passes between a star and the observer; solar – when the ring system passes between the sun and the observer; radio – when a spacecraft broadcasts a narrow radio signal through the rings to a receiver on the Earth. All three are supported by classes which can describe an occultation observation as a time series (how the data is typically captured), or as a derived radial profile of the ring or ring system.

A representative subset of the additional geometric parameters in the rings local dictionary includes a parameter which defines an inertial ring longitude that incorporates the inclination of the ring, a parameter which identifies the extent to which the rings are opened to the observer (from fully open on the illuminated side, through edge on, to fully open on the unilluminated side), a parameter which enables determining the orientation of the observation with respect to the orientation of the self-gravity wakes within the ring, and a parameter which enables determining the location of the observed portion of the ring with respect to the shadow of the parent body projected on the ring.

Since many of the occultation parameters defined in this model also are applicable to atmospheric occultations, a future iteration of the rings local dictionary will include a class or classes supporting atmospheric occultations.

6. The PDS4 application

The PDS4 service structure and its Application Program Interfaces (APIs) are designed to support interactions between disparate systems and make the interactions appear seamless from the end user's perspective, as mentioned above. The PDS4 software is comprised of two main services that support this concept, the Registry Service which is based on the ebXML federated registry model (Organization for the Advancement of Structured Information Standards, 2005) and the Search Service which is based on the Apache Solr (Apache Solr) open source software. The architecture, depicted in Fig. 3, shows metadata harvested from multiple sources into the Registry Service and then indexed and posted to the Search Service.

The Search Service offers a Representational State Transfer (REST)-based interface supporting two protocols for product discovery. The first is a PDS homegrown protocol focusing on common search terms found in the PDS4 Information Model. The second protocol is the International Planetary Data Alliance's Planetary Data Access Protocol (PDAP). The previously mentioned EPN-TAP protocol is also under consideration for support by the Search Service.

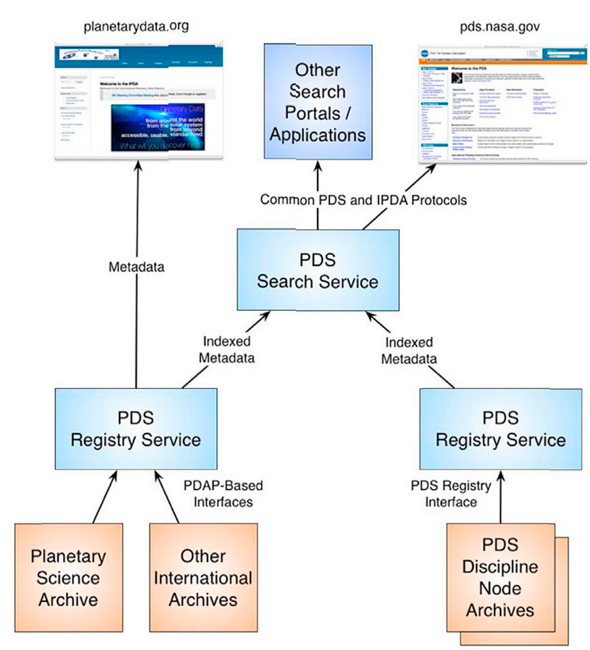

Fig. 3. - PDS4 Service Architecture.

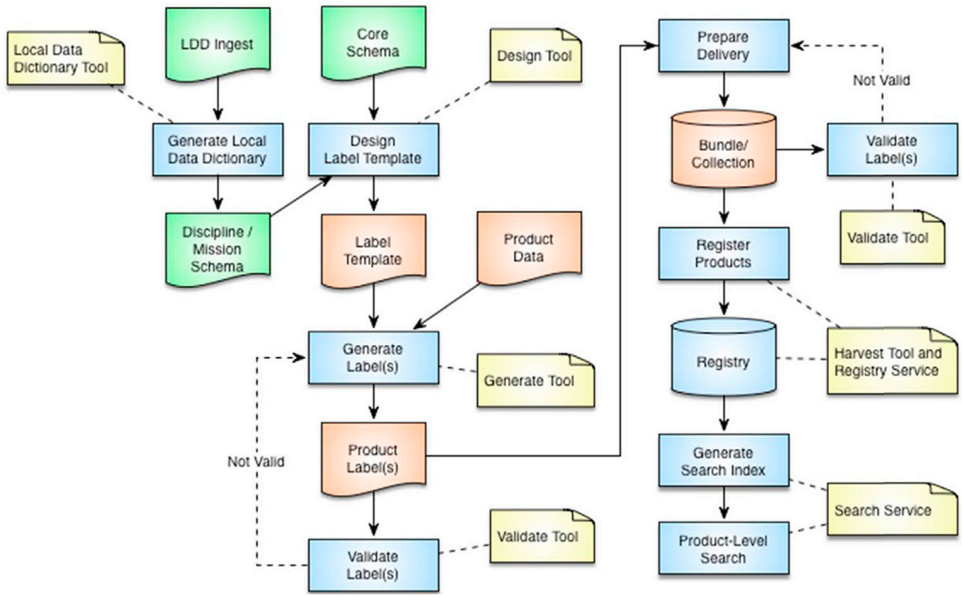

Fig. 4. - PDS4 Information System End-to-End Process.

Finally, the PDS4 end-to-end process is depicted in Fig. 4. As an information model-driven system, the PDS4 Information Model is directly involved in the Generate Local Data Dictionary, Design Label Template, Generate Label(s), and Validate Label Processes. The LDD Ingest, Core Schema, and Discpline/Mission Schemas are all artifacts derived from the Information Model. Extracts from the model are used to configure the tools from the Local Data Dictionary Tool to Product Level Search.

7. Conclusion

The PDS4 Information Model provides the necessary shared knowledge about digital objects in the PDS repository to drive the PDS4 Information System. The PDS4 Information Model is implementation independent and remains relevant within a diverse set of evolving science disciplines.

The Information Model improves interoperability by being ontology-based and having a multi-level governance scheme. The ontology enables interoperability by adding context to vocabulary terms by means of class hierarchies and relationships. Multi-level governance includes a common model for interoperability at the broadest level, discipline level models for science discipline interoperability, and mission models local level interoperability.

In 2002, the National Virtual Observatory Science Definition Team said, "It is probably safe to say that no other professional community has reached the level of data interchange standards (both syntax and semantics) that we have reached in astronomy." (NVO Science Definition Team, 2002) The PDS4 Information Model has enabled a new level of interoperability across the diverse science disciplines of the Planetary Science community. The acceptance of PDS4 by the International Planetary Data Alliance and the planned integration of the EPN-TAP interface and VESPA framework is evidence that the PDS4 Information Architecture is enabling interoperability across the Planetary Sciences and related space science disciplines, such as Heliophysics.

Acknowledgements

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

References

Hughes, J.S., Crichton, D.J., Mattmann, C.A., 2009. Ontology-Based Information Model Development for Science Information Reuse and Integration, 10.1109/IRI.2009.5211603, IEEE International Conference on Information Reuse & Integration.

Hughes, J.S., Crichton, D., Hardman, S., Law, E., Joyner, R., Ramirez, P., 2014. "PDS4: A Model-Driven Planetary Science Data Architecture for Long-Term Preservation", Data Engineering Workshops (ICDEW), 2014 IEEE In: Proceedings of the 30th International Conference on, vol., no., pp.134,141, March 31, 2014-April 4.

Crichton, D., Beebe, R., Hughes, S., Stein, T., Grayzeck, E., PDS4: Developing the Next Generation Planetary Data System, EPSC Abstracts, 6, EPSC-DPS2011-1733, EPSC-DPS Joint Meeting 2011.

ISO 14721, 2003. 2003: Reference Model for an Open Archival Information System (OAIS), ISO.

ISO/IEC 11179, 2008. Information Technology – Metadata registries (MDR), ISO/IEC.

Apache Solr, Apache Lucene, 〈http://lucene.apache.org/solr/〉.

Crichton, D., Hughes, J.S., Sean Hardman, S., Law, E., Beebe, R., Morgan, T., Grayzeck, E., 2014. Scalable Planetary Science Information Architecture for Big Science Data, 2014 IEEE. In: Proceedings of the 10th International Conference on e-Science, 2, 20–24 October 2014.

Besse, S.. et al., 2016. ESA's Planetary Science Archive: Preserve and Present Reliable Scientific Data sets, Preprint submitted for publication to Planetary and Space Science, December 15. this issue.

IPDA Planetary Access Protocol, 2013. V1.0, April 16, 〈https://planetarydata.org/standards/IPDA_PDAP_v1.0.pdf〉.

Erard, S., Cecconi, B., 2014. et. al., The EPN-TAP protocol for the Planetary Science Virtual Observatory, arXiv:1407.5738 [astro-ph.IM].

Erard, S., Cecconi, B., et. al., VESPA: Developing the Planetary Science Virtual Observatory in H2020, EPSC Abstracts, Vol. 10, EPSC2015-270, European Planetary Science Congress 2015.

Bernstein, R., et al., 1982. Data Management and Computation, Volume 1: Issues and Recommendations. National Academy Press.

Arvidson, R.A., et al., 1986. Issues and Recommendations Associated with Distributed Computation and Data Management Systems for the Space Sciences. National Academy Press.

Hughes, J.S., McMahon, The Planetary Data System Distributed Inventory System - An Experimental Interface to a Distributed Scientific Digital Library.

Hanisch, RobertJ., 1999. Distributed Data Systems and Services for Astronomy and the Space Sciences, Space Telescope Science Institute Baltimore, MD ADASS '99, Kohala Coast, Hawaii, 3–6 October.

Hanisch, Robert J., 2002. ISAIA: Interoperable Systems for Archival Information Access. NASA Technical Reports Service (NTRS).

Uschold, M., Gruninger. M., 2004. Ontologies and Semantics for Seamless Connectivity, SIGMOD Record, 33.

Thieman, J.R., et al., 2010. SPASE and the Heliophysics Virtual observatories. Data Sci. J. 9, IGY85–IGY93. https://doi.org/10.2481/dsj.SS_IGY-019.

Hughes, J.S., Crichton, D., Simpson, R., Gordon, M., Joyner, R., Raugh, A., Guinness, E., Martin, M., 2015. The PDS4 Information Model Design Principles - How Well Did They Work?, Proceedings for the 2015 PV Conference, 3–5 November, Darmstadt, Germany.

Organization for the Advancement of Structured Information Standards, 2005. "OASIS/ebXML Registry Information Model V3.0".
Zachman, J.A., 1987. A framework for information systems architecture. IBM Syst. J. 26, 276–292.
Raugh, A.C., Hughes, J.S., 2015. An overview of PDS4 archival file format, astronomical data analysis software and systems XXIV (ADASS XXIV). In: Taylor, A.R., Rosolowsky, E. (Eds.), Proceedings of a Conference Held 5–9 October 2014 at Calgary, Alberta Canada, 2015. Astronomical Society of the Pacific, San Francisco, p. 551.
XML Schema Part 1, 2004. Structures Second Edition, W3C Recommendation, 28 October.
XML Schema Part 2, 2004. Datatypes Second Edition, W3C Recommendation, 28 October.
ISO/IEC 19757-3 Information technology – Document Schema Definition Languages (DSDL) – Part 3: Rule-based validation – Schematron.
Extensible Markup Language (XML), 2008. 1.0 (Fifth Edition), W3C Recommendation, 26 November.
Hughes J.S., Beebe, R., Crichton, D., Grayzeck, E., Huber, L., Neakrase, L., 2012. The Release of the PDS4 Data Standards for First Use by a Planetary Mission, 39th COSPAR Scientific Assembly, Abstract PSB.1-18-12, p. 782.
Content Standard for Digital Geospatial Metadata, 1998. Metadata Ad Hoc Working Group, Federal Geographic Data Committee, FGDC-STD-001-, ⟨https://www.fgdc.gov/standards/projects/metadata/base-metadata/v2_0698.pdf⟩.
NVO Science Definition Team, 2002. Towards the National Virtual Observatory, A Report Prepared by the National Virtual Observatory Science Definition Team.
Lee, Y.T., 1999. Information Modeling from Design to Implementation. National Institute of Standards and Technology.