Developing the International Planetary Data Alliance

Daniel Crichton [(1)], Pedro Osuna [(2)], Maria Teresa Capria[(3)], Reta Beebe[(4)], J. Steven Hughes [(1)], Yukio Yamamoto[(5)], Jesus. Salgado[(2)], Yasamasa Kasaba [(6)]

[(1)] Jet Propulsion Laboratory
4800 Oak Grove Drive, Pasadena, CA, 91011 USA
EMail:{crichton, jshughes}@jpl.nasa.gov

[(2)] ESA-ESAC
Villafranca del Castillo, Spain
EMail:{ Pedro.Osuna, Jesus.Salgado }@rssd.esa.int

[(3)] INAF-ISAF
Rome, Italy
EMail: capria@rm.iasf.cnr.it

[(4)] New Mexico State University
Las Cruces, New Mexico, U.S.A.
EMail:rbeebe@nmsu.edu

[(5)] ISAS/JAXA
Japan
Email: yukio@planeta.sci.isas.jaxa.jp

[(6)] Tohoku University
Miyagi, Japan
EMail: kasaba@pat.gp.tohoku.ac.jp

ABSTRACT

The International Planetary Data Alliance (IPDA) is an international organization with a purpose of developing compatible archives for the capture, management and distribution of planetary science data and results. With the increasing internationalization of planetary science missions, the IPDA is focusing on developing both data and technical standards that support construction of interoperable, online archives across space agencies. Several efforts have been underway since the founding of the IPDA three years ago to develop such standards. First, the IPDA has been working closely with NASA's Planetary Data System (PDS) to leverage PDS data standards. Second, the IPDA has been developing a common protocol for accessing and sharing data as well as defining common information services that can be shared. Third, in order to link the data and technical standards, the IPDA has also been developing a reference information system architecture defines the architectural patterns for both ingestion and distribution, identifying the necessary data and technical standards to ensure that the IPDA can build such an international data resource. This paper will discuss the progress of the IPDA in developing both the standards and the architecture.

Keywords: IPDA, Data Sharing, Planetary Science, Archiving, Interoperability

INTRODUCTION

The International Planetary Data Alliance (IPDA) [2,7,8], an organization founded in November 2006, is an international alliance chartered to develop archiving standards for managing the vast amount of science data results returned from both national and international solar system exploration missions. Over the past decade, the United States and European space agencies have been collaborating on the development and operations of planetary science missions. These collaborations have recently extended to other agencies including Japan and India. Of particular concern is ensuring that data is captured using a common set of data standards and those data are available from public, interoperable planetary science archives.

The IPDA has initiated several projects to develop standards to support interoperability. One of the key projects is to develop a set of data standards for archiving and sharing scientific data products across international agencies and missions. The NASA Planetary Data System (PDS), founded nearly twenty years ago, has developed a set of data standards that are widely used and under adoption by space agencies involved in the IPDA. Since the PDS Standards are considered the "de facto" standards for capturing and archiving planetary science data from solar system exploration, the IPDA and PDS are working to leverage the standards and ensure they are in a form that space agencies can quickly and easily adopt.

A second critical project is to develop a set of technical information system standards allowing interoperability between agency archive systems. These standards are critical to building systems that facilitate users finding, downloading and using data across compatible archives.

Of critical importance to the IPDA is having a system architecture that ties its standards together so agencies can "plug into" and directly leverage IPDA standards with minimal cost. In 2008, an effort was under-taken by the IPDA to develop a reference system architecture that identifies the system and data components as well as the key processes employed to support archive and distribution of planetary science data. The reference architecture serves as a roadmap for developing the standards and defining how the IPDA elements fit together to meet the requirements of the IPDA.

IPDA SYSTEM ARCHITECTURE

The multi-disciplinary nature of planetary science and the increasing number of national space agencies involved in planetary exploration suggest the need for a common system architecture to ease discovery, access and use of planetary data by world-wide scientists regardless of which agency is collecting and distributing the data and to ensure access to and exchange of high quality planetary science data products across international boundaries.

An initial IPDA system architecture has been defined. It consists of three key sub-architectures: a process architecture that describes a set of standard processes for planetary science archive data systems, a data architecture that describes a set of data standards for planetary science archive data systems, and a technology architecture that describes a set of standards for linking planetary science archive data systems. Figure 1 below shows the derived set of elements that make up the IPDA architecture.

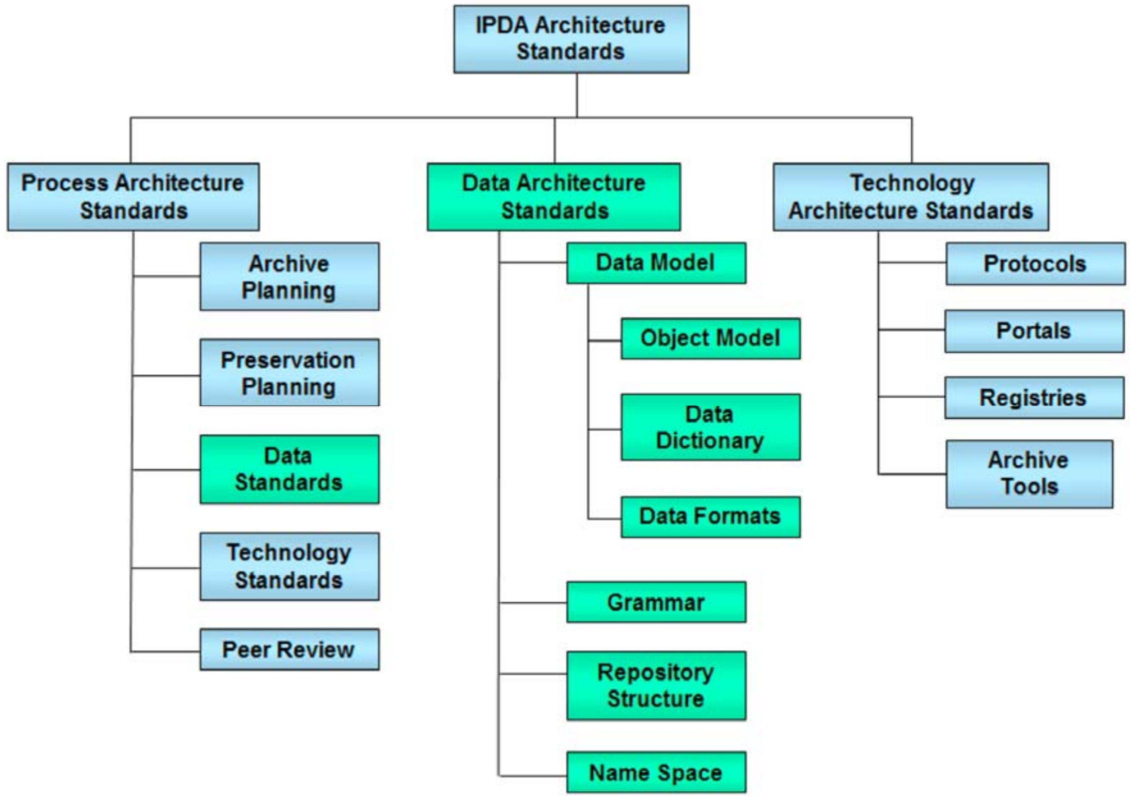

Figure 1: IPDA System Architecture Elements

The following describes each of the sub-architectures from figure 1.

The *Process Architecture* defines the core processes of IPDA (e.g., archive management, preservation planning, peer review, standards management, etc). These elements are critical to the function of an archive system and include elements for the planning, governance and standardization of IPDA-compliant archive systems. Each of the elements is critical to helping IPDA develop compatible archives from the data producer through to management and distribution of the data to the community. Given the number of international missions, coordination of standard processes is important for improving the efficiency in building international archive teams and archiving data in different systems in United States, Europe, India, Japan, etc.

The *Data Architecture* is defined as the structure of an organization's logical and physical data assets and data management resources. The IPDA Data Model is a key component of the data architecture consisting of a data dictionary of terms, standard data formats, and a model of objects and their relationships. Other components of the data standards include a "grammar", or expression of the data model in a particular language (i.e. XML, Object Description Language), repository structure, and processes associated with managing data. While initial efforts have focused on the IPDA Data Model, other planned projects will focus on the other elements which will improve compatibility of archives.

The *Technology Architecture* describes the set of standards and common components for enabling interoperability between planetary science data archive systems and the underlying services. These include standard protocols for searching distributed catalogues of planetary data sets and products, registries for capturing common services and data definitions, and common tools to support generation and validation of data products. In addition, a common portal is provided since IPDA provides and manages this resource in order to provide information regarding IPDA and to link to agency systems.

The initial IPDA system architecture is now under review by the IPDA Technical Experts Group. The intent is to ensure that the architecture and standards can be tied together. In addition, IPDA level 1 and 2 requirements [5], developed and accepted by the IPDA Steering Committee in February 2008, were mapped to the architecture to ensure there were no initial gaps.

DATA STANDARDS

The IPDA Data Standards is the center-piece of the alliance [1]. This provides the foundation for building compatible archives and ensuring world-wide use by the scientific community. The IPDA data architecture provides a set of elements for defining the standards. Critical to this is an explicit data model that provides definition for the structure and content of planetary science archives. Having a coordinated data model is critical as missions become more international and space agencies continue to increase collaboration. To date, the NASA Planetary Data System's (PDS) standards have been adopted and used as the de facto standard for both the agencies and the IPDA. While this has helped to provide a common starting point for building compatible archives, the PDS, along with IPDA, has discovered a need to upgrade the PDS standards and model to ensure that it can be consistently used and adopted across space agencies. In order to achieve this, the IPDA and PDS have laid out the elements of the data architecture and are working very closely to coordinate the future development of planetary science data standards to ensure that the elements are derived from a well-defined architecture and that they can be organized to support both international and local agency needs. Figure 2 shows a planned re-structuring of the standards to partition them for international, national and local use.

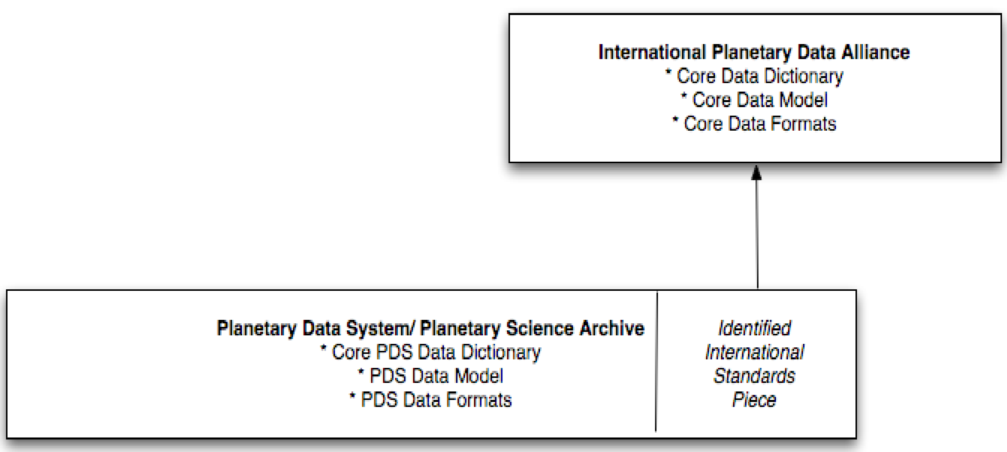

Figure 2: Partitioning of Planetary Science Data Standards

In figure 1, the decomposition of the data architecture includes an element called "name space". As the IPDA and PDS work together to build future data standards [4], one of the principles being applied is to organize the data model into name spaces [5]. This will enable data elements (or keywords) to be categorized in such a way that the scope of responsibility for their management is made at the appropriate level. At the international level, the IPDA can provide a top level set of data elements for annotating data products. At the mission level, for example, which have a need to introduce and annotate a particular data product with a unique data element, can do so by introducing it.

While portioning of the data elements is important to address their governance, the PDS and IPDA have been working to develop a structured data model that represents the planetary science domain. This represents a major upgrade from the existing PDS standards which had a rigorous set of standards, but a loosely defined data model. In the future, planetary science data products should be derived from an explicit data model. One of the major efforts planned for the 2009-2010 period of time for the IPDA, is working with PDS on the definition and upgrade of their data model and standards to improve the consistency and usability and to ensure it can the archiving needs of the international planetary missions.

TECHNICAL STANDARDS

Producing standards to enable sharing data across agency boundaries is one of the critical functions of the IPDA. Figure 3 below shows the relationship between agency data systems and the necessary functional capabilities required to share data. The blue circles in the figure indicate protocols that are required to support data discovery, access and exchange. The IPDA portal, as discussed earlier, serves as an entry point to link to agency data systems as well as disseminate information (standards, events, etc). The IPDA Service and Data Registries refers to registered services and common information that needs to be shared by all member agencies.

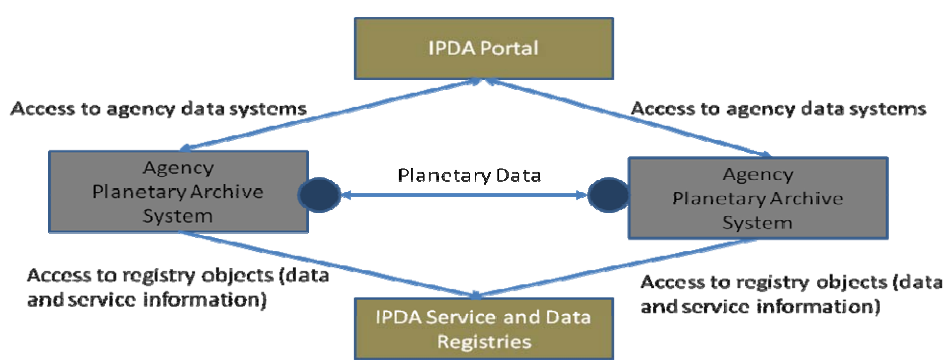

Figure 3: IPDA Technical Architecture

The IPDA has developed the Planetary Data Access Protocol (PDAP). PDAP serves as a standard protocol that describes the attributes necessary for querying for specific data sets and data products across agency systems. PDAP is built using the REST [6] (Representational State Transfer) method to provide simple methods to allow agencies to query across systems. The PDAP attributes are derived from the PDS data standards to ensuring common attributes are used in the construction of queries against the archived and catalogued data set and product metadata. During the 2007-2009 period, the IPDA conducted prototypes using the PDAP specification for Mars Express, Venus Express and the Hayabusa missions. These demonstrated sharing of data between NASA, ESA and JAXA.

In addition, the IPDA has identified the need for common registries. Registries provide an online service to query for information regarding services and data definitions within the IPDA. They also provide standard mechanisms for managing information such as capturing, versioning, updating, replacing and deleting data objects. This includes managing service definitions (physical addresses, descriptions, etc), common data definitions such as mission, instrument information, etc, and other types of data that needs to be standardized and shared across the IPDA.

Finally, the IPDA provides a portal for access to IPDA information and linking to agency systems. Agencies graciously have appointed members to IPDA. In a sense, IPDA is a virtual system. Therefore, little effort is planned for a top-level portal. It is the expectation that agencies will develop their own portals that provide linkages to other systems. The IPDA portal[1] therefore provides links to agency portals as well as information related to IPDA standards.

As the IPDA develops, each of these areas has explicit projects that ensure that development of the technical standards, tools, and services for IPDA will evolve. This will be discussed in the next section.

IPDA PROJECTS AND FUTURE WORK

One of the philosophies of the IPDA is to break development of IPDA into a set of discreet projects. These projects generally run on a yearly basis and are defined at the annual IPDA meeting of the Steering Committee. The Steering Committee appoints a project leader that puts together a plan and recruits members from the agencies.

Several projects are underway in 2009-2010 year. First, as previously mentioned, the IPDA is working with the PDS to plan for version 4.0 of the PDS standards. The intent of the standards is to significantly improve the usability and internationalization of the standards. In addition, the IPDA has initiated a project to assess the existing standards that are in use, and to ensure that they are captured and defined in a central place on the IPDA website. Third, the IPDA is continuing to expand the PDAP specification to support querying of different types of data products. Much of the initial effort focused on a limited set of attributes for image products. Expansion includes not only greater support for querying image products (including leveraging geospatial standards), but also other types of data products. Fourth, the IPDA is developing a plan for definition of common registries in order to support sharing of services and other standard information. In addition, the IPDA has other projects focused on standardizing how navigation and ancillary data is captured and specified, and developing standard processes for archiving.

While IPDA has focused on developing standards, there has also been a concerted effort to perform prototypes in order to both validate the standard and ensure alignment with the agencies. This has helped to ensure that systems, which are being developed in parallel at local agencies, can evolve with the IPDA. It has also helped to emphasize the importance of moving towards international interoperability.

[1] http://planetarydata.org

CONCLUSIONS

The development of the IPDA represents a major step forward in moving towards an integrated, planetary data system to support world-wide scientific research on observational data from planetary science research. The IPDA represents a new era for how those results are ultimately managed and shared. Traditionally, data generated from international missions was captured and managed locally in each agency data system. A new era has begun where agencies now link data through standard protocols. A recent effort between NASA and ESA for Venus Express demonstrated this concept. In this case, agency leaders affirmed the importance by signing a memorandum of understanding (MOU) that emphasized sharing of data through common protocols, rather than managing data locally in both the Planetary Data System and Planetary Science Archive. Through the IPDA, the effort stood up services and shared data using common data and technical standards.

It is the intent of the IPDA to continue to develop and provide standard data, technical and process recommendations for use in agencies. In addition, the IPDA expects to share software tools and experience to ensure leveraging across agencies can help to improve the efficiency and usability of planetary science data and tools. The IPDA is also continuing to work under the auspices of the Committee on Space Research (COSPAR) to ensure that the research community recognizes the IPDA as the de facto organization for developing planetary science data archiving standards. Over its short lifetime, the IPDA has been successful at forming an alliance, working with the research community, developing standards, and coordinating developing across space agencies.

REFERENCES

[1] - D. Crichton: "Core Standards and Implementation of the International Planetary Data Alliance." 37th COSPAR Scientific Assembly, Montreal, Canada (2008)

[2] - D. Crichton, R. Beebe, Y. Kasaba, J. Zender: "Developing the International Planetary Data Alliance". In proceedings of Science Archives in the 21st Century. College Park, Maryland, USA (2007)

[3] – IPDA Steering Committee: "IPDA Level 1 and 2 Requirements". (2008).

[4] – J. Steven Hughes, Dan Crichton, Reta Beebe, Ed Guness, David Heather, Joe Zender: "Defining the Core Archive Data Standards of the International Planetary Data Alliance (IPDA)." PV 2007, Munich, Germany (2007).

[5] - J. S. Hughes, R. Beebe, E. Gunness, M. Huang, Y. Kasaba, P. Osuna, E. Rye, V. Savorskiy: "International Planetary Data Aliiance (IPDA) Information Model". 2nd IPDA Steering Committee Meeting. Pasadena, California, USA (2007)

[6] – Roy Fielding, Richard Taylor: "Principled Design of the Modern Web Architecture". ACM Transactions on Internet Technology (TOIT). New York, New York, USA (2002)

[7] – S. Slavney, R. Beebe, D. Crichton, S. Hughes, J. Zender: "The International Planetary Data Alliance." 38th Lunar and Planetary Science Conference. Houston, Texas. (2007).

[8] – Y. Kasaba, I. Shinohara, J. Zender, R. Beebe, D. Crichton and S. Slavney: "The International Planetary Data Alliance." Asia Oceania Geosciences Society 4th Annual Meeting (2007).