Developing a Core Set of Data Standards for the IPDA

CONCEPT WHITE PAPER

January 2007

I. Introduction

"This decade has witnessed a rapid expansion of Solar System exploration. Two major, parallel trends can be observed. Space agencies fly smaller, short-lived dedicated spacecraft as well as some larger spacecraft hosting suites of very complex instruments. And, where in the past, the planetary exploration was the domain of NASA and the Russians, then joined by the Europeans with ESA, it is obvious that today more countries and space agencies are capable of and are demonstrating their ability and commitment to be involved in space exploration, e.g. CNSA (China), JAXA (Japan) and ISRO (India). Often spacecraft observations are complemented by observations from earth-orbiting and ground-based facilities."[1]

In November 2006, the 1st meeting of the International Planetary Data Alliance was held at the ESA Technical and Engineering Center (ESTEC). Several space agencies attended the meeting including NASA/PDS, ESA/PSA, JAXA/ISAS, CNSA and RSA/IKI. The meeting helped to define needs and constraints of the international planetary science data community including both a need for a set of core data standards for building planetary science data archives as well as the budgetary constraints that necessitate leveraging of standards and tools across space agencies. The ultimate goal of the IPDA is to ensure that agencies can share and reuse science data products.

II. Compatibility vs Compliance

The IPDA affirmed the importance of developing a set of core standards that can be used to form the foundation of the IPDA data standards. In particular, the IPDA confirmed the need to seek IPDA compliance over IPDA compatibility. Compatibility requires a transformation occur between two systems in order to share data since each system is built from its own set of governing standards. Compliance ensures that no transformation or mapping is required in order to use a data product between two systems. Effectively, compliant data products generated by any member archive would be readily accessible to users of other agencies even though the physical structure in member archives will be determined by local constraints. This concept is important in understanding the future development of standards for the IPDA. The selection of the "compliance" model over the "compatibility" model means that agencies are committed to building planetary science data products from

[1] NASA/ESA Planetary Data Interoperability White Paper

one core set of data standards. Figure 1 shows the transformation between two compatible domains that require a transformation in order to mediate between both domains.

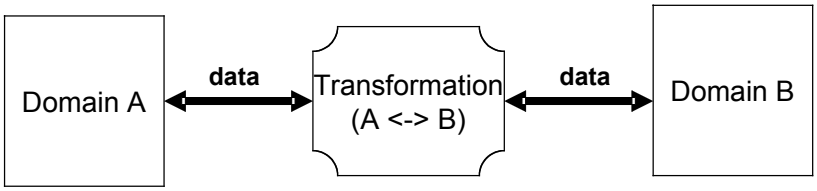

Figure 1: An example of compatibility

As the IPDA seeks a core, it recognizes the importance of allowing extensions to the core to support agency or local needs.. The intent of the IPDA is to provide a core standards set on which agency specific data standards can be built in order to enable IPDA-level compliance across agencies. Figure 2 shows an example of the relationship between the PDS data standards and the IPDA data standards.

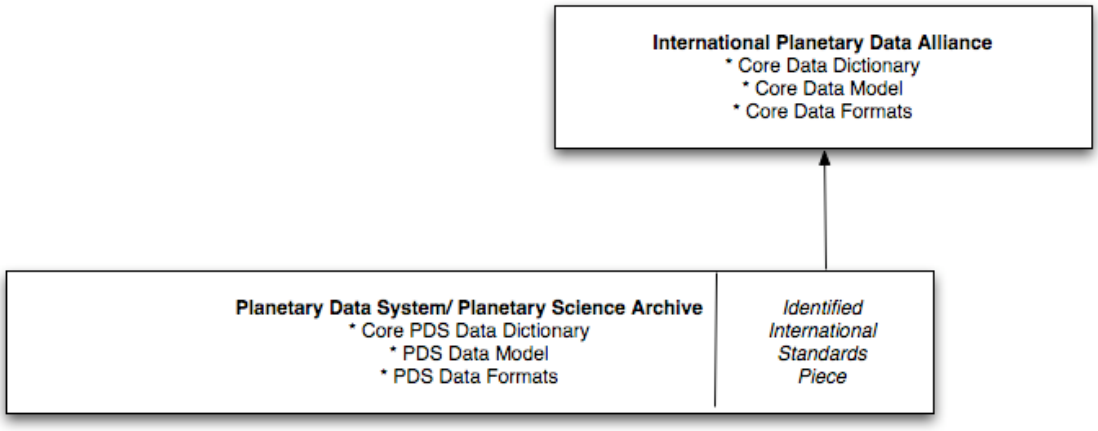

Figure 2: Example IPDA-to-Agency Data Standards Relationship

III. Elements of the core

An IPDA Reference System Architecture has been developed that consists of three key elements, 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. The key component of the IPDA Archive Data Standards is the IPDA Data Model consisting of a Data Dictionary of Terms, Standard Data Formats, and a Model of

Objects and their Relationships as shown in Figure 3. Other components of the Data Standards include a "grammar", or expression of the data model in a particular language (i.e. ODL), repository structure, and processes associated with managing data.

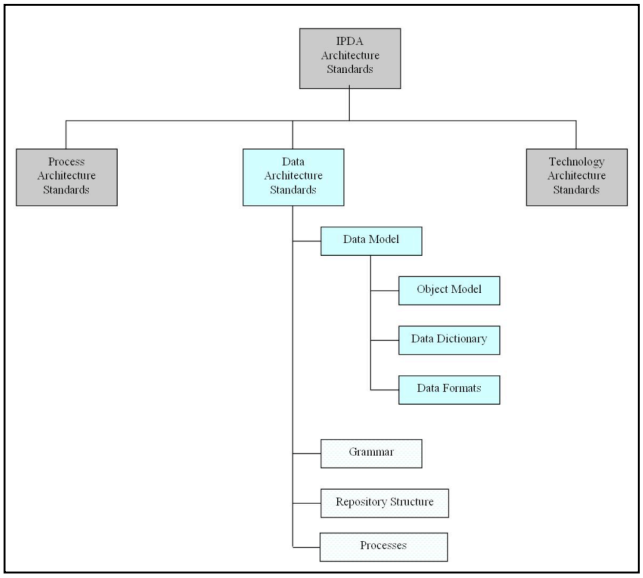

Figure 3: IPDA Architecture Standards

IV. Engineering process to get to the core

At the November 2006 meeting, the IPDA recommended that the Planetary Data System data standards be used as a basis for generating the core. Specific motivations for this included the use of the PDS data standards by the European Space Agency's Planetary Science Archive as well as PDS' vast experience supporting archiving for the planetary science community and an established international community of users. Of particular concern to agencies which are beginning to fly planetary missions, is the cost of building such a system from scratch. The IPDA therefore recognized that it was important to generate a "core" that builds on PDS experience and development.

As previously mentioned, the core is divided into multiple components including:

- Core data model consisting of an object model, data dictionary and set of data formats
- Grammar for describing and documenting planetary science data
- Organizational structure for documenting the data holdings
- Process for managing the standards

Derivation of the core data standards is decomposed into four separate projects including:

Project 1: Generation of the IPDA data model
Project 2: Generation of the IPDA grammar
Project 3: Generation of the IPDA standard archive processes
Project 4: Generation of the IPDA archive organizational structure

Each of the above deliverables will follow a process to generate requirements and then a core recommendation for the standard.

Project 1 to define the IPDA data model is the first project the Steering Committee has recognized. Efforts are underway to define the requirements. Given that the IPDA has decided to use the PDS Standards as a basis for the IPDA standards, the PDS data model is being formally documented as an initial step. In addition, the PDS standards adopted by ESA's Planetary Science Archive will be used to help identify common terms, formats and an object model. This process will also help identify problems with the existing models that can then be addressed. A survey of data sets and data formats at other member agencies will also be done to help determine the core elements of the data model. Project 1 is intended to wrap up by July 2007 in order to support the 2nd IPDA Steering Committee Meeting in Pasadena, California July 17-21.

V. Project Approach

The IPDA has adopted a project-centric approach to managing its development. The IPDA Steering Committee identifies an issue or need and then forms a project with a team lead to address it (ie. Project 1. Define the core requirements and document the data model for the IPDA).

The suggested approach for managing an IPDA project requires that the project leader identify a schedule and accept the responsibility of frequent and direct communication with the IPDA Chair and individual members of the committee. Management and operation of an IPDA Project shall be based on the following steps.

1. The Chair of the Steering Committee (with the help of the committee) identifies and clarifies the project.
2. The Chair appoints a Project Leader and (in consultation with members of the steering committee) establishes an expected duration duding which the project should be completed.
3. The Project Leader (in consultation with the Chair) develops an outline, including a series of time steps, for accomplishing the task and picks a team.
4. The Project Leader circulates the outline to the Steering Committee for input, approval and suggestions (including justification and identification of skills) for additional team members.
5. The Project Leader will guide a two-step process. Beginning with the first time step

a. The Project Leader poses a set of questions and assembles material to address this component of the project and specifies a time interval for initial input.
b. At the end of the specified interval the Project Leader integrates input from the Steering Committee and places that component of the document on a web page, where the committee members can access it and supply comments to the Project Leader. The leader can modify the posted version as he/she deems appropriate.

6. The Project Leader then repeats steps 5a & b for the additional project time steps that have been accepted in the plan.
7. After all project steps have been developed, edited and presented to the Steering Committee for comment, the Project Leader and members of the team will generate a final document and submit it to the IPDA Chair for acceptance and development of implementation plans.

VI. Timeline

The first of the four projects associated with the IPDA Archive Data Standards is underway with representations from NASA, ESA, JAXA/ISAS, RSA/IKI, and CNSA including R. Beebe, E. Guinness, D. Heather, M. Huang, Y. Kasaba, P. Osuna, E. Rye, V. Savorskiy, and S. Hughes. The project team is currently reviewing a set of use cases and preparing an initial set of requirements. Concurrently the team will be identifying PDS data formats either currently in use by the PSA or being considered for use by other agencies such as JAXA and the SELENE project. These will be included in an initial data model to be presented at the 2nd IPDA Steering Committee meeting in July.

Other projects which are currently being defined are as follows:

Project 2: Generation of the IPDA grammar
Project 3: Generation of the IPDA standard archive processes
Project 4: Generation of the IPDA archive organizational structure

These projects will be defined at the IPDA Steering Committee meeting in July and will follow the same structure as Project 1. In addition, NASA and ESA will do initial preparation for the July 2007 meeting in order to document the current "As-Is" standards for each these areas within the PDS and PSA.