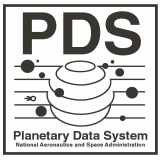
# NASA-PDS/ESA-PSA Planetary Data Interoperability

NASA/ESA

White Paper

By Pedro Osuna, Steve Hughes, Joe Zender and the NASA-PDS/ESA-PSA teams


July 2005

# Contents

Section 1

# 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 Russian Space Agency, then joined by ESA, it is obvious that today more 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.

Recent ESA and NASA successes have defined the breadth and volume of data that will be generated in the future and growing interest in and plans for lunar and Mars exploration form the structure for an International Lunar Decade and expansion of the exploration of Mars. In contrast, the difficulty of reaching other solar system bodies imposes limited temporal coverage and generates datasets that must be integrated to obtain maximum yield. Although significant international collaborations have been established, members of the planetary community need inter-agency data access to realize optimal use of the data.

This White Paper addresses a solution for locating and retrieving scientific data across Space Agency boundaries and from differently structured data systems.

A small team of engineers and scientists from the NASA Planetary Data System (PDS) and the ESA Planetary Science Archive (PSA) has considered how to implement an agency independent data query and retrieval standard. As a first step towards such an ambitious goal, this White Paper proposes the adoption of a limited standard protocol to demonstrate interoperability between PDS and PSA to provide reciprocal access to their data.

A major challenge in developing a worldwide Planetary Science Interoperability system will be the multi-disciplinary nature of planetary science. It is clear that the overall activity for developing an international standard must evolve over time and be supported by more space agencies and scientific groups. Initially, we are proposing to adopt a simple standard protocol. Using this protocol, we will demonstrate the viability of interoperating NASA PDS and ESA PSA data at a high level by building a client prototype application that will consume this protocol. This effort will scope the problem and allow us to develop more precise and powerful protocols for expanding capabilities and allowing additional agencies to consider joining the effort.

This White paper aims to provide input to the respective project-level managers responsible for developing road maps and high-level requirements for PSA and

PDS. The following sections contain the goals, advantages, effects on the existing systems and an outline of the protocol.

## Goal

A group of engineers and scientists from PDS and PSA proposes to approach the aforementioned interoperability issue according to the following two-step process:

1. The definition and adoption of a Planetary Data Access Protocol restricted to mapped, gridded planetary surface data.
2. The implementation of a simple proof-of-concept prototype supporting query and retrieval functionality in both directions.

The prototype shall support data from ESA's High-Resolution Stereo Camera (HRSC) flown on the Mars Express spacecraft and NASA's Gamma Ray Spectrometer (GRS) flown on board of the Mars Odyssey spacecraft.

This approach will demonstrate that:

- query and retrieval of data from Space Agency archival systems can be seamlessly integrated into existing interfaces;
- the implementation only depends on the draft protocol standard (and not on the exchange of information between engineers);
- the prototype implementation is cheap and quick (less than three months, excluding the definition and agreement of the Protocol standard).

## Current Status, Impact on Archival Systems and long term vision

The European Space Agency (ESA) developed the Planetary Science Archive (PSA) that was released in early 2003. Today, the PSA contains 0.5 TBytes of ground-based data and data from the Giotto and Mars Express spacecraft. The

data from all ESA spacecraft will be ingested into the PSA, as e.g. Huygens, Smart1, Venus Express and Rosetta. ESA requests all Principal Investigators to create and deliver data following the PDS Standard (currently 3.6). To ensure the long-term support and cost effectiveness of the PSA, ESA shares the core of the archive functionality (development, maintenance, system) between all planetary and astronomy missions. As such, the interoperability functionality is already available from existing astronomy efforts, see IVOA in the references section, as it is inherited from previous astronomy archives, e.g. the XMM-Newton and ISO archives (see http://esavo.esa.int/). The development of the proposed prototype and also further achievements will be seamlessly integrated into the existing astronomy and planetary archival system of ESA and have no negative impacts. Currently, ESA solar and Earth plasma missions are not included in the common archival approach.

The Planetary Data System (PDS) is the official science data archive for NASA's planetary science community and currently holds about 20 terabytes of data collected from over thirty years of solar system exploration. Although the archive is geographically distributed across several science discipline nodes and other locations, it provides a single point of entry web service for retrieval of metadata and data. The PDS data system integrates the distributed archive repositories and provides search and retrieval capability for data products across the majority of the data holdings. The Object Oriented Data Technology (OODT) connects the individual Discipline Nodes. With this proven technology, also here the proposed prototype will seamlessly integrate in the existing system and no negative impacts are foreseen.

Further implementations of the Planetary Science Interoperability Standards would have an important impact on both archival systems. With a given, accepted protocol standard, any other, third party can develop an interface with both archives. Future developments could include:

1. Enhanced indexes that incorporate observational pointing and geometry to support gated searches.

2. Powerful graphical user interfaces that allow discipline related queries and retrievals across all collaborating space agencies.

3. Data analysis packages that data mine all the archival systems, giving the user selected attributes and allowing limited, efficient data transfer.

4. Query translators, attribute mediators, reference frame converters and data format converters to meet the common standard.

The implementation of these might require considerable investments on the long-term.

We are entering a new era in data distribution and analysis and the real impact on the archival systems will be manifold, but are difficult to anticipate.

Section 2

# Standard Protocol

## The proposal

As a first step towards the definition of a Planetary Science Interoperability System, the PSA and PDS propose to define a high level Planetary Data Access Protocol (PDAP) for data access.

Both PSA and PDS will create servers at their respective ends implementing the agreed protocol, i.e., understanding protocol conformant queries and responding with protocol conformant answers.

The system will provide the users, the international science community, with the capability to access (query, select and retrieve) science quality data products, independent of the data provider or the location of the data products.

It is proposed that the initial system will be accessed both through http (for human readable access) and through raw sockets (for machine access).

The basic functionality of the protocol will consist of a two-step process:

1. The server (either PSA or PDS) is queried to get the metadata corresponding to matches with respect to certain input parameters

The server responds with metadata corresponding to matches (including pointers to the real data)

2. The server is queried to deliver the real data for some of the inputs.

The server delivers the data

The basic system input/output functionality advantages reside in the translation from different entities to common standardized ones:

1. Syntactically varying "query strings" are converted to a "standard query string"

2. Semantically varying "keywords" (that might depend on specific projects) are converted to a set of "standardized" keywords

3. Syntactically varying xml based output (and prone to project dependent specifications) are converted to "standardized" VOTable (xml-like) output.

Examples of valid instances of the protocol follow for reference:

**http://psa01.sciops.esa.int/PAIO/pdap.jsp?TARGET_NAME=MARS&INSTRUMENT_NAME=HRSC**

**(Valid instance of http metadata request)**

```
<?xml version="1.0" ?>  <!DOCTYPE VOTABLE SYSTEM "http://us-vo.org/xml/VOTable.dtd">
<![CDATA
[http://psa01/?DATA_SET_ID=MEXMHRSC[...]&PRODUCT_ID=H0521_0000_S12PG]
]>
      <TD>MEX-M-HRSC/SRC-3-RDR-0501-V1.0</TD>
      <TD>H0521_0000_S12.JPG</TD>
  </VOTABLE>
```

**((Highly simplified) valid metadata standardized output. Note the pointer to the real data in the CDATA section)**

The technical details of the proposed protocol can be seen in reference [1].

## The implications

The goal of this task is to develop a prototype to show simple interoperability between the PSA and the PDS data systems while causing minimal impact on the existing infrastructures. Impact on the existing infrastructure can occur in three areas, the data system queries, domain data models, and the formats of results.

### *Impact on Query (NASA-PDS)*

The prototype will have minimal impact on the existing PDS query mechanisms. The chosen standard is in line with the ESA Archive Inter-Operability (AIO) protocol. The PDS engineering staff studied documentation on the AIO protocol and find it to be very similar to the PDS-D query protocol. A PDS mediating interface will be developed to translate between the two query protocols and that presents the AIO query protocol to external systems. This mediating interface will translate incoming AIO queries into equivalent PDS-D queries and vice versa for outgoing queries.

### *Impact on Data Model (NASA-PDS)*

Science domain experts from the PDS and PSA have compiled a set of data elements that can be used to search for data products in both data systems consistently. Since the PSA and PDS data dictionaries share a common source, only minor differences exist between data elements. Where differences between a PSA and PDS data element do exist, for example different data element names, the PDS mediating interface will make the necessary translations.

### *Impact on Results Formats (NASA-PDS)*

The PDS returns metadata results in an XML structure called a profile. The PSA returns metadata results in an XML structure called a VOTable. PDS engineering staff believes that the PDS mediating interface will be able to easily convert the VOTable into an XML profile for incoming results and vice versa for outgoing results.

### *Impact on existing Infrastructures (NASA-PDS)*

The PDS staff feels that the PSA/PDS interoperability prototype will have little impact on existing PDS infrastructure since a mediating interface is planned to perform the translations necessary for both incoming and outgoing queries and results. Any changes to the PDS model, namely the modification or addition of data elements to the Planetary Science Data Dictionary, will be handled in the mediating interface.

### *Overall Impact on ESA-PSA group*

The proposal follows very closely the efforts done by the IVOA in defining standard protocols for astronomical data access. ESA group involved in the PSA and AIO design is also involved heavily in the definition of the aforementioned IVOA protocols and standards. The systems designed by this group, including the AIO system, are quite in line with the interoperability philosophy and therefore it is not foreseen that there will be major impacts on development strategies to implement the proposed protocol.

Moreover, the PSA is fully designed to be compatible with our NASA colleagues' PDS, and therefore, the integration with NASA PDS dictionary, etc., should not impose a major burden in the road map.

In summary, no major impacts besides those normal of engineering efforts are foreseen on existing systems for the implementation of the proposed prototype.

### *Overall Impact*

The development of the long-term vision, the Planetary Science Interoperability System, will proceed in two phases. Phase one, as described in this proposal, is the definition of the PSA/PDS Data Access Protocol and the construction of the proof-of-concept prototype client that will make use of the NASA and ESA protocol-aware servers.

There are already several existing agreements regarding the definition of the protocol to be adopted, and the example data sets. Also it is expected that the existing data system infrastructures can easily be adapted to meet the requirements of this proof-of-concept system. The resources required for implementation are estimated to be about 4 working weeks for each of the two participating data systems.

The implementation of the basic interoperability functionality of the PSA is foreseen for September 2005. Based on this system, it is expected that an additional 10 days will be required to tune to the specific protocol requirements. Implementation and testing will be conducted by ESA Sciences Archives and VO team. An additional 5 days are foreseen to review documentation and execute the required testing.

For the PDS it is expected that five days of support will be required across the team members for planning, document review, and subsequent testing of the proof-of-concept. It is expected that five days of support will be required from the Geosciences Node for the completion of the few tasks remaining to make a PDS data set accessible under the OODT, namely refinement of the Data Access Protocol and the installation of a profile server. It is expected that about 10 days support will be required from the Engineering node for software development and task management.

Phase two of the Planetary Science Interoperability System, the implementation of a more comprehensive inter-operation capability, will start only after a thorough technical and management review of the proof-of-concept system, testing results, and assessment of characteristics of existing data sets. It will be decided at that time whether multiple development phases will be required.

Section 3

# Conclusion

This White Paper describes the task to develop a simple prototype to demonstrate the capability to search and retrieve science data products across the PSA and PDS data systems. A standard protocol that addresses query, data modeling, and data formatting aspects of the interaction will be adopted for the prototype. Each participating data system will then use mediating interfaces as necessary to interact using the standard protocol. This prototype will have minimal impact on the existing data systems and will provide a proof-of-concept for consideration of continued development of a more comprehensive interoperable data systems interface.

Section 4

# References

[1] NASA-ESA Planetary Data Interoperability

(Draft Proposal from ESA-Planetary Science Archive, SOP-RSSD-TN-028, 2005-04-01)

http://www.rssd.esa.int/SYS/include/pubs_display.php?project=PSA&id=477754

[2] International Virtual Observatory Alliance http://www.ivoa.net/

Section 5

# Acknowledgements

The authors would like to thank the ESA-PSA and NASA-PDS teams:

***ESA-PSA***: John Dowson, Inaki Ortiz, Aurele Venet, Jose Hernandez, Guillermo San Miguel, Jesus Salgado, Isa Barbarisi, Christophe Arviset, O. Witasse, D. Heather, J. Diaz del Rio, Joe Zender, Pedro Osuna

***NASA-PDS***: Mike A'Hearn, Reta Beebe, Dan Crichton, Ed Guinness, Steve Hughes, Ron Joyner, Sean Kelly, Susan Slavney, and Thomas Stein