

IPDA Assessment Results Build 1c

Steve Hughes

August 2011

Topics

- Introduction
- Phase 1 – Assessment
- Phase 2 – Prototype
- Classification of Responses
- Next Steps

Introduction

- At the International Planetary Data Alliance (IPDA) Steering Committee Meeting held July 2009 in Rome, Italy, the IPDA Steering Committee felt that it was important to initiate a project to prototype agency data sets using the new PDS4 Data Standards¹.
- The goal of the project is to test the capabilities of the PDS4 Data Standards for preparing data sets for use in the international planetary science community under the auspices of the IPDA.

¹The PDS4 design team also wanted external input from our partners.

Project Team

- P. Allan
- M. Gangloff
- D. Heather
- G. Krishna
- T. Roatsch
- A. Sarkissian
- I. Shinohara
- T. Stein
- M. Teresa
- J. Salgado
- S. Hughes
- M. Gordon
- R. Joyner

Two Phases

- Phase 1 - Review¹ the PDS4 Data Standards and documentation and complete an assessment response sheet.
 - Start April 15, 2011
 - End May 15, 2011
- Phase 2 - Prototype PDS4 data product using one or more data products produced by the agency.
 - Start May 25, 2011
 - End² July 15, 2011
- Both phases used the PDS4 Data Standards documents that had been released for build 1c.

¹The IPDA also checked that their recommendations from the 2009 IPDA assessment were addressed.

²D. Heather and S. Martinez's prototype report is promised soon.

Phase 1

Assessment Questions

1. Do the document provide sufficient background for the review? If not, how could they be improved?
2. Assess the four fundamental structures. Are they useful? Will they support your needs? Do you have products that you believe will not fit into the structures?
3. Assess the PDS4 core product types. Do they provide an adequate set of baseline templates for constructing new templates and new PDS4 products? What is missing?
4. Assess the structure and layout of the PDS4 product examples? How can it be improved?
5. What overall recommendations do you have for the team? Do you have suggestions for improvement?

Assessment Response Summary

Results: 157 Response Items

- Overall, I found them to be well written and to be fit for their intended purposes. It is a big step forwards from what was available previous. - *P. Allan in general*
- In general (it) gives a good description of an archive structure but consistency issues and some clarifications are needed ... - *D. Heather and S. Martinez on Stds Ref*
- There are a lot of TBD, but the structure of the document and the general contents is appropriate. - *M. Gangloff on Stds Ref*

Question 1

Do the document provide sufficient background for the review? If not, how could they be improved?

Sample Responses

1. Yes, most of important points are indicated at this point and level of achievement indicated. I should propose to add some more details relevant to nodes at this step.
2. Yes, but a lot of paragraphs are still missing in some documents
3. When read as a whole, the set of documents under review appear to contain enough information to understand the fundamentals of PDS4 and, in combination with the examples, provide the appropriate background to start the design and preparation of a PDS4 archive. Nevertheless, more accurate descriptions, requirements, rules and recommendations are expected to be provided in the Standards Reference to better guide the data providers in the correct use of the PDS4 Standards.

Question 2

Assess the four fundamental structures. Are they useful? Will they support your needs? Do you have products that you believe will not fit into the structures?

Sample Responses:

1. I had difficulty do [sic] understand what you call four fundamental structures here.
2. Yes, most products are table based and array based.
3. Fundamental structures are useful and easy to understand. They seem to cover most of our (PSA) needs, although we haven't found any way to:
 - (a) Describe an ISIS3 qube with Tile format ...

Question 3

Assess the PDS4 core product types. Do they provide an adequate set of baseline templates for constructing new templates and new PDS4 products? What is missing?

Sample Responses:

1. Yes, they are good templates. And the capability to define extensions for specific missions and nodes is a good idea.
2. Phase2 of this assessment will allow us to better answer this question. In order to identify missing product types, we need to start using PDS4

Question 4

Assess the structure and layout of the PDS4 product examples? How can it be improved?

Sample Responses:

1. The structure and layout of the product examples are easy to read. They are a good help when trying to understand the PDS4 concepts.
2. Some errors, inconsistencies and areas of confusion found in the examples have been reported in "PDS4 Issues".

Question 5

What overall recommendations do you have for the team?
Do you have suggestions for improvement?

Sample Responses:

1. The XML introduction and terminology is very nice and useful to those of us with limited XML experience, but consider to include this as an Appendix with a reference in the Introduction.
2. Is there really any need for two Dictionary versions (abridged and unabridged)? So far, we haven't used the abridged version at all.

Phase 2

Prototype Questions

1. How well did the process for creating PDS4 products work?
 - Is the generic product schema you chose complete and useable?
 - Is the process for creating a specific schema well documented and complete?
 - Is the process for creating a product label well documented and complete?
 - What parts of the process could be improved or what needs to be changed?
2. What tools should be developed and made available?
3. Are the PDS4 data standard documents useful?
 - What could be improved?
4. Did you find any limitations or items missing that you expected?
5. Do you have any other comments?

Prototype Response Summary

Results: 31 Response Items

- The creation of PDS4 labels is straightforward for people with a minimum knowledge of XML and XML Schemas - *M. Gangloff*
- As people start to use PDS4, to what extent is it possible for two people working independently to produce product labels for the same product that are inconsistent? - *P. Allan*
- I would have found it useful if the documentation included a "toy" example that was deliberately cut down in the amount of information provided - *P. Allan*

Question 1

How well did the process for creating PDS4 products work?

- Is the generic product schema you chose complete and useable?
- Is the process for creating a specific schema well documented and complete?
- Is the process for creating a product label well documented and complete?
- What parts of the process could be improved or what needs to be changed?

Sample Responses

1. The Template schema chosen (Product_Table_Character_0311B.xsd and Product_Table_Binary_0311B.xsd) are useable
2. The generic product schema is useable. Their [sic] can be more tags related to the spacecraft orientation and altitude
3. For the local identifier reference more clarity [sic] is required.

Question 2

What tools should be developed and made available?

Sample Responses:

1. A tool that generates web interfaces from a schema for label generation.
2. A tool to convert existing PDS3 datasets to PDS4 Bundles
3. A tool for reading/viewing the PDS4 dataests [sic]
4. A search engine to search a specific [sic] product from the bundle with some criteria of the orbit number or start_date_time

Question 3

Are the PDS4 data standard documents useful?

- What could be improved?

Sample Responses:

1. PDS4 data standards documents are very useful for this exercise. I also used som [sic] examples of schemas and labels
2. They are useful and appear to be comprehensive.

Question 4

Did you find any limitations or items missing that you expected?

Sample Responses:

1. I did not get serious problems, but I did not have enough time to test all schemas.

Question 5

Do you have any other comments?

Sample Responses:

1. Data providers not familiar with XML , XML schemas and Oxygen will need some help from discipline nodes and/or engineering Node
2. Please comment on the attached label generated for the downloaded datasets [sic]

Classified Assessment and Prototype Responses

Ambiguous	- 6
Duplication	- 2
Error	- 10
Improvement	- 3
Incomplete	- 22
Inconsistent	- 21
Missing	- 19
Other	- 39
Question	- 4
Suggestion	- 41
Kudos	- 21

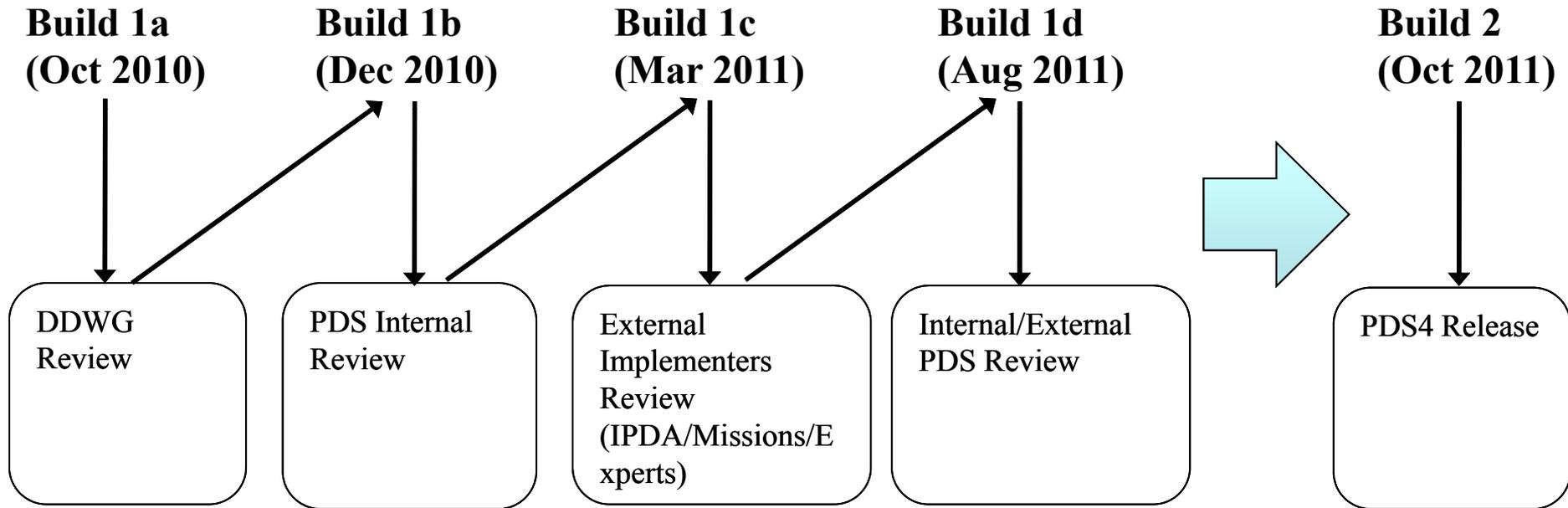
Classified Response Examples

1. Ambiguous - What would be used as Product Class if combining more than one class in one product?
2. Duplication - The Concepts and DPH documents contain extensive descriptions in some sections
3. Error - Table_Character example with fields, records, record_bytes=0 should not be allowed.
4. Improvement - update document bundle or collection
5. Incomplete - Chapter 4 is a good introduction to labels , but it must be completed.
6. Inconsistent - basic products are called standard products
7. Missing - add a list of current namespaces used
8. Other - I do not understand what 'inter-process communication' means in this context.
9. Question - Why should I start now to learn how to use XML editors, how to write a PDS4 image file by hand, etc. ?
10. Suggestion - I suggest to add a « tutorial » containing an actual example of a bundle delivery

Next Steps

- The document editors have the assessment comments and are currently updating their documents.
- The prototype comments are being distributed.
 - D. Heather and S. Martinez's are promised soon.
- Snapshots of all PDS4 Standards documents will be released for build 1d on 8/29.
 - Information Model and schemas will be frozen on 8/22.
 - Build 1c issues will be resolved by Build 2.

PDS4 Assessment/Input Process



Links

- Results are posted to the IPDA PDS4 Assessment and Prototyping Wiki Site
 - <https://oodt.jpl.nasa.gov/wiki/display/pdscollaboration/IPDA+PDS4+Assessment+and+Prototyping%2C++April+-+July+2011>

Backup

Questions

Phase 1

Significant Results - 1

- In general the assessment is that good work is being done and that the PDS4 data standards will be useful to data providers in the Planetary Science Community. Some of the more broad recommendations are:
 - Better separation of the Concepts and DPH documents so that the DPH contains just the basic guidelines and requirements for data preparers/users and with the most common practices/examples.
 - For IPDA, it would be nice to see references to how PDS4 should be used for data providers not delivering directly to a PDS Node. This could be a simple matter of generalizing by adding something like 'your Archiving Authority' to the sections describing management of schemas and dictionaries.
 - The XML introduction and terminology is very nice and useful to those with limited XML experience, but suggest that this is an Appendix.

Phase 1

Significant Results - 2

- Consider eliminating the abridged data dictionary since it was not used in the assessment.
- The Standards document needs quite a lot of work. This will be the central resource for data providers wishing to use PDS4, so requirements need to be very clear in this document, and recommendations only used for non-critical elements.
- Clarification of Data dictionary management procedures including rules for editing/modifying the product types, the link between the XML schema and the Dictionary, formation rules or permissible values from the dictionary are in the schema, management of schema the Node / Archive Authority level.
- Continue toward the use of more intuitive language especially for the sake of non-native English speakers.

Phase 2

Significant Results

- In general the assessment is that good work is being done and that the PDS4 data standards will be useful to data providers in the Planetary Science Community. Some of the more broad recommendations are:
 - Better separation of the Concepts and DPH documents so that the DPH contains just the basic guidelines and requirements for data preparers/users and with the most common practices/examples.
 - For IPDA, it would be nice to see references to how PDS4 should be used for data providers not delivering directly to a PDS Node. This could be a simple matter of generalizing by adding something like 'your Archiving Authority' to the sections describing management of schemas and dictionaries.
 - The XML introduction and terminology is very nice and useful to those with limited XML experience, but suggest that this is an Appendix.