December 05, 2013

To: PDS4 ORR Review Board

From: Dan Crichton, PDS Engineering Node Manager

Re: Disposition of the PDS4 Board Report and RFA

On behalf of the PDS4 project, please find attached the disposition of the PDS4 ORR RFAs. We would like to express appreciation and thank the Panel Members for their participation, open discussions and insight during the review. The comments and recommendations are thoughtful, highly helpful and constructive.

In addition, the PDS4 project is providing the build 4a test documents and requirements traceability in order to address the lien against the deployment of the software and information model.   The build 4a software testing has been completed and pending the board review of the test results, we plan to begin deployment to prepare for the LADEE and MAVEN missions.

Please let us know if you have any questions.

December 05, 2013

Response to PDS4 Operational Readiness Review Panel Report

Background

PDS4 Operational Readiness Review

The Planetary Data System 4 (PDS4) Operational Readiness Review was held September 17-19, 2013 at the NASA Goddard Space Flight Center. The Panel membership was:

Brian P. Duncan, JHU/APL, Chair

PJ Clark, JHU/APL

Dave Heather, ESA/ESAC

Dave Linick, JPL (via telecon)

Tom McGlynn, GSFC

Lauren McIver, Virginia Bioinformatics Institute (via telecon)

Marilyn Newhouse, MSFC/CSC

Peter Shames, JPL

The intent of the System Review was to assess the readiness of a limited release of PDS4 for data providers for the LADEE and MAVEN missions.

The review panel report was received on November 06, 2013. The report included an Executive Summary, Assessment Against Review Objectives and a set of Requests for Action (14).

PDS4 Project Response

The PDS4 Project would like to express appreciation and thank the Panel Members for their participation, open discussions and insight during the review. The comments and recommendations are thoughtful, highly helpful and constructive. The PDS4 Project provides this response by disposing Request For Action (RFAs), addressing the Panel's concerns and recommendations, and answering Panel Members' questions.

The Panel has submitted 14 formal RFAs. Detailed disposition information for these RFAs is provided below.

Title: PDS4 ORR	RFA: PDS4ORR-RFA1	Category: Lien	Date: Sept 17-19. 2013

Author: Brian Duncan, JHU/APL; Marilyn Newhouse, MSFC/CSC

Statement of Concern: It is difficult to understand the testing performed (and results generated) to validate and verify the core tools to support LADEE and MAVEN.

Explanation: The scope specified in the ToR for this ORR includes the "Test plan and results for this Limited Release". The entrance criteria specified in the ToR indicates that the test plan and results will be made available to the review participants in advance. The assumption of the review panel going into the ORR was that the testing of the system for LADEE and MAVEN was complete and documented. Through discussion during the ORR it was understood that testing was on-going, development of the information model and software was on-going: release 3b is more or less closed out, but release 4a is preparing for the test cycle and is needed/desired for LADEE/MAVEN. To support common data format (CDF) files and data product-level searches, development and testing must continue for the next revision of the information model and the 4a release of the software. For the review panel to gain confidence that the testing approach is sufficiently mature that a stable 3b system would not be jeopardized by changes made in 4a, the panel dug into the testing performed on 3b, the requirements traceability, and the artifacts generated as a result of the testing. From this research, it was not easy to tell what testing had been done, the details of the system under test, and the results of that testing.

Recommendation: The review panel requires that some modifications to the capture and presentation of test material be adopted during the 4a test cycle. If the following are applied to the 4a test cycle and documented, then this lien will be addressed for closure.

a. A new requirements trace to the test case has been supplied, but we were still not able to tell pass/fail for individual requirements, only for overall test cases. In the test <u>results</u> report, please indicate the pass/fail of each requirement covered by that test case. If a test case fails, for example, it does not necessarily mean that all of the requirements failed.
b. In the test results report, it is important to indicate the version under test. If the test result report shows the effect of testing multiple internal software builds over time, then this is most easily handled by indicating the software build version number in the test case result section (where you already have the date).
c. In the anomalies section (i.e. JIRA) of the test results report, we would like to see the data more clearly organized and with some further information:
    a. Separate the list of issues so that it is clear which were submitted during this round of testing. We should know, for example, which issues were submitted new during testing of release 4a (versus ones that were submitted back in 3b, etc.)
    b. Separate (or at least show in the table) the issue type so that it is clear what is a software defect, what is a software improvement, etc.
    c. Please show in the table the severity of each issue; a list of 5 minor defects may not be as significant as one showstopper defect
    d. Please show in the table the title of the issue, so that the reader can get a sense of the

types of issues being found during test

d. An executive summary at the end of the test result report showing a requirements verification matrix, and what has passed or failed for each requirement provides the health of the release at a glance. The Build 3b system test plan very nicely shows the requirements and test case IDs that test each (in section 4). This table with a test **result** for each would give us a succinct view of the state of the release at a glance.

Response:

The PDS4 team has adopted and implemented the Panel's requirements of the capture and presentation of test material for the Build 4a test cycle. This includes:

1. Updated requirements traceability matrix

2. Full testing of the software and information model

3. Full documentation of captured test results

4. Updated and reorganized test documentation as follows:

   o Combined the system test plan, test procedure and test report into a single document to ease maintenance of test case/trace matrix

   o Combined the Node test plan, test procedure and test report into a single document to ease maintenance of test case/trace matrix

   o For each test case, specified what requirements were tested and showed individual requirement had been passed or failed during testing

   o Identified CCB changes to XML schema and schematron since build 3b and included results of testing them

   o Added version numbers for software components to tell what version of software was under test o For JIRA list of issues, indicated clearly what issues were found during what build testing, and for each issue whether it's a defect or an improvement o Included in the traceability matrix pass/fail status

The Build 4a test and requirements traceability documentation will be made available to the Panel for review.

Title: PDS4 ORR	RFA: PDS4ORR-RFA2	Category: Action	Date: Sept 17-19. 2013

Author: Peter Shames, JPL

Statement of Concern: The document relationships specified in the LADEE LDEX ICD and the SIS are not clear. The ICD references other SIS documents that may or may not exist.

Recommendation: Clarify the ICD document in the areas where it refers to a SIS. Is this as simple as only referring to a single SIS that now exists? Perhaps provide a template for an ICD and the necessary SIS documents to support it. This should clarify what is expected of other data providers.

Response:

The following is a clarification for the board regarding ICD and SIS documents that are used by PDS for agreements with the missions and instrument teams.

The ICD is the document that:

    a. Defines the responsibilities of the Mission and the PDS
    b. The product is the MISSION BUNDLE - The Mission Bundle contains the Mission description (mission.cat) the spacecraft description (insthost.cat) and possibly useful mission planning and operation.

The SIS is the document that:

    a. Defines the responsibilities of the individual instrument team and the PDS archiving node.
    b. The product is the INDIVIDUAL INSTRUMENT BUNDLE.

Section 1.4 of the LADEE LDEX ICD specifies the "Relationships with Other Interfaces". The ICD specifies responsibilities for writing LADEE LDEX data product SIS.
Both LADEE LDEX ICD and SIS are available on the ORR web site. In addition, example ICD and SIS can be found on PDS home website: http://pds.nasa.gov/tools/archiving.shtml#examples which provide templates for the nodes and missions in developing agreements.

Title: PDS4 ORR	RFA: PDS4ORR-RFA3	Category: Action	Date: Sept 17-19, 2013

Author: Dave Heather, ESA/ESAC

Statement of Concern: The current / proposed set up for MAVEN and LADEE focuses on web-based tools, software and some documentation external to the bundles being archived in order to provide the best user service. This is fine for currently active missions but requires significant maintenance and implies much information and many tools may be lost if the www system is no longer managed. This goes against the long term usability of the data.

Recommendation: Ensure that all information (documents, algorithms, calibrations etc) from the web pages that may be useful to end users is included within the bundle itself, so that data can be used and understood fully with the standalone bundle should the full online support be discontinued. This could become a standard check made as part of the PDS4 Peer Review process.

Response:

After a mission is over, PDF documents and text files carrying titles such as accessing the data can be constructed and deposited into the individual instrument bundles. These files would be generated from the online web pages to capture as part of the permanent archive. This would assure that information such as indexes generated to access the data were preserved over the long haul. We are considering an internal certification step to assure all information is included in the final bundles.

Title: PDS4 ORR	RFA: PDS4ORR-RFA4	Category: Recommendation	Date: Sept 17-19, 2013

Author: Peter Shames, JPL

Statement of Concern: The overall handling of requirements in the project could have benefited from some of the same level of rigor that has been applied, with great effect, to the information model.

Recommendation: Use of any formalism (DOORS, SysML, or even a more traditional and complete spreadsheet) would have been a better way to provide traceability down to design and to the test suite. Given the present information there is no strong evidence that all requirements have been satisfied.

Response:

The PDS4 team will maintain the requirements and traceability as part of a requirements spreadsheet. This has been integrated with the Build 4a test results and will be maintained by PDS going forward.

Title: PDS4 ORR	RFA: PDS4ORR-RFA5	Category: Recommendation	Date: Sept 17-19, 2013

Author: Lauren McIver, Virginia Bioinformatics Institute

Statement of Concern: It would be good to clarify what a "browse" collection is in the Standards Reference (pg 9) and the Concepts document (pg 8).

Recommendation: I would suggest adding examples to these documents as was done with the other collection types in the Standards Reference (pg 10). Another option would be to include references in the Standards Reference and the Concepts document to the definition found in the Data Providers handbook (pg 24).

Response:

The following parenthetical text will be added to page 9 of the Standards Reference:

"A *browse* collection contains 'quick-look' products (such as low-resolution JPEG images or plots of table data) designed to facilitate use of the archive. Products in a browse collection are defined using appropriate classes; possibilities include Product_Browse and Product_Thumbnail."

The following parenthetical text will be added to bullet on page 8 of the Concepts Document:

"Browse products (such as low-resolution thumbnail images)"

In addition, wording to explain the purpose of a "browse" collection will be included in the Concepts Document.

Title: PDS4 ORR	RFA: PDS4ORR-RFA6	Category: Recommendation	Date: Sept 17-19, 2013

Author: Dave Linick, JPL; Peter Shames, JPL

Statement of Concern: While the PDS is a "federation" of nodes, and while recognizing that the needs of the nodes vary one from another, the PDS4 architecture offers opportunities to more fully exploit commonality among the nodes and to better integrate the PDS into a true system.

Recommendation: The PDS project would really benefit from having a higher level of top down architecting & engineering to guide and focus the Engineering Node and Discipline Nodes for common look and feel of web sites, common/adaptable tools, and tool sharing.

As a recommendation for Program Management: Develop a PDS vision and a Strategic Plan for achieving that vision. The vision should be bold, recognize the full possibilities of the PDS, and exploit the PDS4 architecture to its fullest. It should include a roadmap that defines, at a top-level, how the Strategic Plan should be achieved over time. The strategic Plan should be owned by the primary stakeholders, including the Program, the nodes, and the science community that uses the PDS. This would serve as an alignment vehicle and would provide a basis for proposing a logical and defensible evolution of the PDS. Much progress has been made recently and this is an opportunity to build on that progress.

Response:

This RFA has been delivered to the PDS Program Management for consideration in updating the current PDS roadmap by the PDS Management Council.

Title: PDS4 ORR	RFA: PDS4ORR-RFA7	Category: Recommendation	Date: Sept 17-19, 2013

Author: Lauren McIver, Virginia Bioinformatics Institute

Statement of Concern: The term "node" has multiple definitions.

Recommendation: In the concepts document, section "1.2 The archiving process" and "7.3 Validation" it seems a "consulting node" and "discipline node" are PDS staff. In the same document, section "6.4 Using data dictionaries" it refers to "node data dictionaries" that are "needed by users of that node for searching or manipulating the data" in which a node seems to refer to a group of resources. Since "consulting/discipline node" seems to be a commonly used term throughout PDS, it is probably not possible to change "node" to "staff" for instances in which it refers to an individual. Instead it might be useful to define the term in the Concepts document when it is first used. I would also suggest adding the definition of "discipline node" to the PDS4 Glossary at the end of the Concepts document as currently a "consulting node" is defined as a "discipline node" but there is no definition for a "discipline node" or "node". The definition could state it is a group of resources that include both PDS staff and data sets. This would help in clarifying these terms for individuals who are not familiar with the PDS terminology.

Response:

A footnote at the beginning of the Concepts Document (and in the Glossary) will be added as follows:

(a) The PDS team will add to the Concept Document a footnote after the first sentence in section 1 to read:

"Discipline nodes are the research centers of PDS and set the priorities for PDS, curate the discipline-specific data, conduct continuing research on those data, and provide access to the data for the entire planetary science community. A discipline node also provides consultation to users of those data, participates in technical development of PDS, and develops and disseminates archive standards. A discipline node may be headquartered at a single institution or distributed over several, some of which may be known as 'subnodes' or 'data nodes'. (NRA 03-OSS-04)"

The PDS team will also add the following entry to the Glossary:

"discipline node – an entity having local governance within the federated Planetary Data System. The following are currently recognized discipline nodes: atmospheres, engineering, geosciences, imaging, management, Navigation and Ancillary Information Facility, planetary plasma interactions, rings, and small bodies. Discipline nodes curate discipline-specific data, make them available to users, provide consultation, and participate in technical development and management of the PDS.

(b) When the current version of the Concepts Document was written, there was a distinction between

'node data dictionaries' and 'mission data dictionaries'. Those are now consolidated into 'discipline data dictionaries.' When the documents are revised, this change will be incorporated.

Title: PDS4 ORR	RFA: PDS4ORR-RFA8	Category: Recommendation	Date: Sept 17-19, 2013

Author: PJ Clark (JHU/APL); Dave Heather (ESA/ESAC)

Statement of Concern: The PDS documentation should have sections clearly describing restrictions placed on external file standards such as FITS and CDF that will allow them to be made fully PDS4-compliant.

Recommendation: A specific section should be included in the Standards Reference and/or the Data Provider's Handbook to list the restrictions that are placed on all acceptable external file standards. This should cover both CDF and FITS as needed for MAVEN. Ensure that the information contained in the backup slide entitled "Requirements for Archivable CDF" becomes a registry document at any discipline node that archives CDF files. Ensure that all labels that pertain to CDF encoded products contain a reference to this document in the labels' Reference_List.

Response:

PDS understands the concern. However, PDS can only enforce its own standards; it cannot tell data providers which modifications to their external formats will make them PDS compliant because PDS does not have that expertise. However, PDS can provide suggestions on how those formats could be tailored in documents that are not part of PDS 'standards'.

The PDS Management Council has discussed this issue and will provide information about how data in a select set of formats (e.g., FITS, CDF, etc) has been structured for inclusion in the PDS4 archive.

Title: PDS4 ORR	RFA: PDS4ORR-RFA9	Category: Recommendation	Date: Sept 17-19, 2013

Author: Dave Heather, ESA/ESAC

Statement of Concern: If a restricted form of CDF is going to be an accepted format for MAVEN data, and accepted by PDS in general, the tools, software and documentation developed by PPI specifically to accommodate this should be centralized or endorsed by the PDS as a whole so other future nodes using the format will not have to develop their own tools for the same purpose.

Recommendation: Centralize the tools and documentation being developed by PPI for the usage of CDF within PDS4. At least the PDS, at a high level, should endorse them and ensure that the same set of tools is used for future CDF deliveries to the PDS.

Response:

PDS will provide a registry of commonly used tools for sharing. This will include tools developed by the PPI node for CDF if CDF is accepted by PDS. It will also be shared with the IPDA.

Title: PDS4 ORR	**RFA:** PDS4ORR-RFA10	**Category:** Action	**Date:** Sept 17-19, 2013

Author: Lauren McIver, Virginia Bioinformatics Institute

Statement of Concern: Running some of the PDS tools on a windows platform I came across a section in the generate, validate, and transform tools code that would benefit from a warning message.

Recommendation: Line 26 of the windows batch file for these tools sets the JAVA_HOME environment variable to a fake path (set JAVA_HOME=\path\to\java\home). If JAVA_HOME is not defined in the environment and the user does not alter line 26 of the code, the code fails with an error message (a default message printed when the system tries to execute at a non-existing path) that is not the best description of the issue for the failure. This message is "The system cannot find the path specified." I would suggest putting in a warning message at line 26 instead of setting the variable to a fake path. This warning can direct the user to the readme which describes how the user should set the path. This would make the code more user-friendly.

Response:

The PDS4 team will make the above suggested modification to the UNIX scripts and Windows batch files. Since this situation applies to all of the scripts in the PDS4 software tree, the modifications will not be complete until the Build 4b delivery.

Title: PDS4 ORR	RFA: PDS4ORR-RFA11	Category: Recommendation	Date: Sept 17-19, 2013

Author: Dave Heather, ESA/ESAC

Statement of Concern: It is unclear if the data to be delivered by LADEE and MAVEN are going to flow through the full validation cycle or if it will evolve significantly from what was shown during this review, and therefore it is difficult to judge whether the system is ready now to allow full ingestion and product level searches.

Recommendation: PDS could require data providers to run the full validation tool as it evolves to check bundles and collections. Wherever possible the validation tools and procedures should be centralized.

Response:

PDS does not have a policy that requires a data provider to run a specific tool prior to delivery of data to the Discipline Nodes even though PDS does specify requirements for valid data products, collections and bundles in the Standards Reference. And, PDS does provide the Validate Tool to the PDS community and promotes its use by Data Providers in their pipelines. This RFA has been delivered to the PDS Program Management and the PDS Management Council for consideration in developing a policy on what PDS expects data providers to do in terms of validation.

With respect to the functionality of the Validate Tool, it will continue to evolve with each build, adding functionality and refining the aspects of validating PDS data submissions.

Title: PDS4 ORR	RFA: PDS4ORR-RFA12	Category: Recommendation	Date: Sept 17-19, 2013

Author: Marilyn Newhouse, MSFC/CSC

Statement of Concern: The goals for the PDS4 development, while laudable, are not presented as metrics that can be used to prove the completion and success of the PDS4 development.

Recommendation: The PDS4 goals (for example, see slide 12 of the presentation) should be quantified, and the progress against these goals measured. This provides several benefits: 1) a clear definition of when the PDS4 has realized its goals and is "complete, " 2) measurement of progress and early identification of concepts or approaches that may be straying from the desired path, and 3) data useable by NASA HQ to justify and defend PDS4 progress and funding. The metrics should look at the PDS holistically, that is, from a perspective that includes development, maintenance, and mission effort and cost to ensure that emphasizing improvements in one area do not simply shift effort to another area. Some historic data may need to be estimated at this point, but defining the metrics by which the PDS4 improvements will be measured will identify what data should be collected in the future. Some possible metrics, provided mainly to begin a brainstorming process, are:

Improve efficiency and support to deliver high quality science products to PDS	- Effort required (PDS & mission personnel) to define the products and interface for mission data delivery (should decrease) - Effort required (PDS & mission personnel) to deliver and validate the first and subsequent data deliveries (should decrease) - Number of redeliveries required to provide validate mission data (should decrease) - Percent budget allocated to maintenance (should decrease) versus development of new capabilities (should increase)
Improve user support and usability of the data in the archive	- Total number of data requests (should increase) - Total number publications using archived data (should increase) - Amount of data retrieved vs. session duration (should increase) - Number of requests for assistance in finding and retrieving data (should decrease)

Response:

The PDS will develop specific metrics for measuring the efficiency of PDS systems in order to quantify the effectiveness of the implementations and to identify areas of improvement in system efficiency. These will be coordinated with the PDS Management Council.

Title: PDS4 ORR	**RFA:** PDS4ORR-RFA13	**Category:** Comment	**Date:** Sept 17-19, 2013

Author: Tom McGlynn, GSFC

Statement of Concern: The PDS4 Software Tools and Services for LADEE/MAVEN describe a capability in the Search Interface of: Provides the second line of metadata harvesting within the system in order to facilitate discovery of products. If this is a future expansion capability to enhance the search metadata, we would suggest that it be controlled in a similar manner to other software design/implementation changes.

Response:

The PDS4 team accepts the recommendation and although the capability is often presented as a limitless mechanism for introducing additional metadata to the Search Service, the options currently under consideration are following the PDS4 design process. The main option utilizes the Product_Update product for incorporating updated or missing metadata for data products at the time of search index generation.

Title: PDS4 ORR	RFA: PDS4ORR-RFA14	Category: Comment	Date: Sept 17-19, 2013

Author: Dave Linick, JPL

Statement of Concern: APPS, the Automated Pipeline Processing Service from AMMOS, is likely to be available in time to support MAVEN data ingestion into PDS. There was no mention of this possibility in the ORR. This is not inappropriate, since APPS is not a pre-condition for operational readiness for MAVEN, but APPS should be considered in the eventual plan for MAVEN.

Response:

This comment was forwarded to JPL Mission Interface Office (MIO) for consideration. The MIO has been coordinating all effort with the MAVEN Mission Manager and generated a signed Service Level Agreement (SLA) and Statement of Work via the JPL Contracts Office.