July 27, 2010

Response to PDS 2010 System Review Board Report

Background

PDS 2010 System Review

The Planetary Data System (PDS) System Review was held on March 22nd, 23rd, and 24th, 2010 at the JPL Offices in Washington, D.C. The Board membership was:

David Heather, ESA, ESTEC
David Korsemeyer, ARC
Dave Linick, JPL, Board Chair
Jan Merka, GSFC
Andy Schain, NASA HQ
Peter Shames, JPL

The intent of the System Review was to assess the architecture and design of the next generation of the PDS (PDS-4) and to assure that the implementation plans, and the deployment and transition planning is adequate and appropriate.

The final review board report was received on June 02, 2010. The report included an Executive Summary, Assessment Against Review Objectives, Individual Board Member Comments, a set of Requests for Action (20), and the Summary Debriefing to the PDS management Council.

PDS 2010 Project Response

PDS 2010 Project would like to express appreciation and thank the board members for their participation, open discussions and insight during the review. The comments and recommendations are highly helpful and constructive.

PDS 2010 Project provides this response by:

- Disposing Request For Action (RFAs)
- Addressing the Board's concerns and recommendations
- Answering Board Members' questions

The board has submitted 20 formal RFAs. Detailed disposition information for these RFAs can be found in Attachment 1. A summary is listed in Attachment 2. Responses to comments provided by board member Dave Korsemeyer are provided in Attachment 3. Responses to comments provided by board member Andy Schain are provided in Attachment 4. Responses to concerns and recommendations listed in the Executive Summary section of the board report are as follows.

Responses to concerns and recommendations listed in the Executive Summary

1. **The engineering rigor applied to the implementation of PDS 2010 should be better specified and appropriate to the task.**

An appropriately tailored NPR-7120.8 is being applied to the PDS 2010 Project including review process. A mapping of NPR-7128.8 requirements and PDS 2010 Project Plan is included in Attachment 5.

2. **Better define the balance between the centralized and decentralized approach.**

PDS 2010 Project is mindful in striking a balance between the benefits of having autonomous discipline nodes and the benefits of a centralized approach. It is made specific through the definition of governance processes and system-level requirements directed by the PDS Management Council (MC), including requirements that define system-wide standards, including those related to common look-and-feel, and that define common tools. The system allows for changing the balance based on the directions and policy given by the PDS MC.

3. **The design and management of the information models and data dictionaries needs to be better defined.**

The design of the information models includes a specification that is clear and unambiguous. The process that manages the information models and their changes is being carefully considered. It will be documented and delivered with the final specification. A well-defined data dictionary based on ISO-11179 is in place.

4. **Clarify the priority of and resources allocated to the international component.**

The design of the PDS 2010 information model and system does not preclude future international collaboration/integration. In addition, resources outside of PDS 2010 effort are allocated within PDS to support IPDA and enable international interoperability.

5. Clearly specify when, and under what conditions, missions and new data providers must be compliant with the new PDS-4 format.

This concern is under considered by the PDS Management Council.

Attachment 1

RFA Disposition Details

Title:	PDS 2010 System Review	RFA:	1	Date:	22-24 Mar 2010
Author:	Jan Merka			Email:	jan.merka@nasa.gov
Topic:	The core data model and data dictionary are crucial for distributed queries				

Comment/Concern: The data model and dictionary are critical components of the PDS4 architecture but their development is not much discussed.

The data model is planned to consist of 'core' and 'local/extended' parts where the latter will be created by DNs or specific data domains. The content of the core data model will however determine the type and features of distributed queries possible within the PDS system. The model development and management plans unfortunately lack sufficient details how this will be achieved.

Recommendation:

- The model development should be fast-paced and firmly controlled to allow its quick employment.
- The model should be comprehensive and well tested with strict core rules.
- The extended/local models should be directed by templates and/or examples provided by the core model developers.

Response:

PDS agrees with the recommendations. The model development is fast-paced and firmly controlled to allow its quick employment. The final model will be comprehensive and well tested with strict core rules. Work is currently being done by the Data Design Working Group to ensure that the extended/local models will be directed by templates and/or examples provided by the core model developers.

Status:

Closed

Title:	PDS 2010 System Review	FRA:	2	Date:	22-24 Mar 2010
Author:	Jan Merka			Email:	jan.merka@nasa.gov
Topic:	Data node holdings overlap in content				

Comment/Concern: It is unclear to what extent data node holdings overlap in content and how is this managed.

The PDS data nodes archive products but sometimes a product might be (for convenience or performance reasons) available from several data nodes. The presentations did not touch on this issue, so it is unclear whether this situations in fact occurs at PDS and if yes how is it handled. In particular, how would this affect the archiving process and/or the distributed searches (cross-DN searches).

Recommendation:

- Please provide a brief discussion addressing my concerns.

Response:

Yes, this situation does occur and will most likely continue under PDS 2010. The Registry Service supports the concept of replicating a registry entry allowing one Node to replicate a product registration from another Node's instance of the Registry Service. The curating Node maintains ownership of the registry entry and control of any updates to the entry. In addition, a given product registration may have more than one physical (on disk) instance associated with the registration. In this case, one instance will reside at the curating Node and one would reside at the Node with the replicated registration.

This allows local searches (local to each Node) to return the registered product in the result set and provide a link to the local instance of the product files. A cross-Node search performed against the aggregated instance of the Registry Service will return one result for the product and provide a link to the curating Node's instance of the product files.

The current archiving process dictates that the curating Node is responsible for archiving the product, which will continue in PDS 2010. The other Nodes simply have copies of the product and are not responsible for archiving.

Status:

Closed

Title:	PDS 2010 System Review	RFA:	3	Date:	22-24 Mar 2010
Author:	Jan Merka			Email:	jan.merka@nasa.gov
Topic:	How much NASA funding goes towards supporting collaboration with international partners?				

Comment/Concern: PDS2010 is a NASA-funded project that needs to work closely with international partners. The budget however does not specify what portion of funding goes towards this task.

International collaboration is extremely important for PDS2010 functionally and, ultimately, for its success. The PDS budget should address, at least estimate, how much effort is dedicated to supporting international collaboration.

Recommendation:

- Estimate how much effort will go towards supporting the international components and parts of PDS2010

Response:

Resource has been allocated in each node's POP budget outside of the PDS 2010 project to provide international support.

It is also noted by the board discussion that there is a great deal of interest in international "interoperability." The PDS should implement the architecture so as not to preclude this, but needs to remember that this is a NASA effort. Nothing should be done to enable interoperability that detracts from meeting the requirements to satisfy the NASA community. This is a principle reason IPDA was started.

Status:

Closed

Title:	PDS 2010 System Review	RFA:	4	Date:	22-24 Mar 2010
Author:	Jan Merka			Email:	jan.merka@nasa.gov
Topic:	What is PDS2010 relationship to the NASA Virtual Observatories?				

Comment/Concern: The NASA Virtual Observatories (VxOs) have been developing distributed searches driven by SPASE metadata for several years and their work may be relevant to PDS2010.

It was not clear to me whether the PDS team considered technologies and approaches used by the VxOs. They may not be applicable for the PDS2010 environment but a brief discussion would be appropriate. Several question about the potential PDS-VxO relationship come to mind, for example: Will the VxOs point to or search the PDS data? Will PDS leverage any VxO services? Will there be a translation of the PDS data model to/from the SPASE data model?

Recommendation:

- Please address the PDS-VxO interactions, if any.
- If the VxO are irrelevant for PDS2010, briefly comment on the reasons.

Response:

PDS did look at SPASE and is interested in adopting some of the SPASE data model. There won't be any direct link among PDS, VxO and IVOA. Todd King and Steve Joy are members of SPASE team and they participate in PDS 2010 system and data design working groups and influence PDS 2010 technology and approach decisions based on their SPASE experience and involvement. In addition, IVOA, VxO and IPDA are all discussing potential collaborations.

Status:

Closed

Title:	PDS 2010 Review	RFA:	5	Date:	22-24 March 2010
Author:	Dave Heather			Email:	dheather@rssd.esa.int
Topic:	Documentation Inconsistencies and comments				

Comment/Concern:

- Harvest Tool SRD/SDD section 8.0, and Registry Service SRD/SDD section 8.0: '…no planned phasing with regard to the implementation with all planned capabilities available in Build 1.' Does the Registry Service not rely on input from the Harvest Tool? There must be some sort of phasing for these in order to work.
- Report Service SRD/SDD: I don't understand or see the difference between Use cases described in Sections 3.3. and 3.4. Please clarify.
- PDS 2010 Operations Concept: Section 4.3 mentions a 'Registry/Inventory Service' and a 'Harvest/Ingest service'. Should be plain 'Registry Service' and probably just 'Harvest service'.
- PDS2010 Operations Concept: Section 5.2.17. The first point here states that the 'Data Provider assembles a peer-review committee'. Should it not be the PDS Node that does this?

Recommendation: Clarify and update documentation as necessary

Response:

Design and development of the Harvest Tool and Registry Service are occurring concurrently with both components scheduled for deployment in Build I. Although we do anticipate change requests and further development of these components in future builds, the first build will contain the core capabilities defined by the requirements.

For the Report Service, use case 3.3 (Create Report) focuses on specifying the criteria for the report or report template, which defines the content of the report. Use case 3.4 (Export Report) focuses on the report format (e.g., HTML, PDF, etc.).

In the Operations Concept, the reference to Registry/Inventory pertains to the instance of the Registry Service containing inventory-related items (e.g., data products, mission descriptions, etc.). This is consistent with the later references in the document to the Document Service instance of the Registry Service. The "Harvest/Ingest services" reference, is a reference to two distinct components for PDS 2010. The Ingest Service is targeted for Build III. That said, section 4.3 has been reworded for clarification.

In the Operations Concept, the reference to "Data Provider" is erroneous. The document has been updated accordingly.

Status:

Closed

Title:	PDS 2010 Review	RFA:	6	Date:	22-24 Mar 2010
Author:	Peter Shames			Email:	peter.shames@jpl.nasa.gov
Topic:	Support for global data searches (#1 intro, #16 architecture, and elsewhere)				

Comment/Concern: ***The capabilities for global and targeted data searches have yet to be developed.***

The team has quite rightly deferred working on this until some of the seminal data modeling and data standardization work has been done and until the ingest functionality has been defined. Given available resources this is entirely understandable. This work seems to be a next priority.

Recommendation: My recommendations:

- As part of the design process identify next, in detail, how the global and targeted searches work within the architecture
- Ensure that global searches are adequately supported while also supporting the different types of targeted and local searches that have been identified

Response:

PDS agrees with the recommendations and that has been the intent. Design of the Search Service is underway with deployment planned for Build II.

Status:

Closed

Title:	PDS 2010 Review	RFA:	7	Date:	22-24 March 2010
Author:	Dave Heather			Email:	dheather@rssd.esa.int
Topic:	Harvest and Registry Tool Comments				

Comment/Concern: Just a few comments that may need clarification in the documentation.

The Harvest tool requires cfg file input by Discipline Node, but it is not clear to me why this should be needed. The purpose of the tool is to crawl through the data holdings and identify new or updated products. Why not automatically scan for new items / time stamps in the data holdings using an xml image to map the last run of the tool?

It is not clear from the documentation if the possibility will exist for the Harvest and Registry tools to handle the situation where there is a need to step back to an older, previously ingested version of a data product or data set. For example, if a data provider releases a newly calibrated product, and later finds serious issues with the calibration, they may request to remove access to the latest file and to make the previous version available. I assume this could work with the 'Deprecate Artifact' Registry function, but it was not 100% clear from the documentation that this had been considered.

Recommendation: Clarify the points in the documentation.

Response:

The configuration file is available for the Harvest Tool to specify settings for a given execution of the tool including any settings that may be specific to the executing Node or environment. Its existence and use does not preclude the mode of operation described above (as detailed in use case 3.2).

In response to the second point, that is exactly the use for the Deprecate function of the Registry Service. The Registry Service offers the general actions of Approve, Deprecate and Delete. It is up to PDS policy to define when these actions are appropriate.

Status:

Closed

Title:	PDS 2010 Review	RFA:	8	Date:	22-24 Mar 2010
Author:	Peter Shames			Email:	peter.shames@jpl.nasa.gov
Topic:	Improve architecture description and understanding (#17 Service Design)				

Comment/Concern: ***#17, pg 7 is very confusing, looks like layered chart but it really just a list of standards related to client, presentation, and logic layer clients and servers, needs a better approach such as RASDS or SCA***

Most of the architecture discussion (aside from data architecture) is focused on software architecture, little truly addresses system architecture. This chart is an example. While it looks like a representation of a layered architecture it is little more than some notional representation of layers along with a set of standards that might be relevant to the layer. In order to make clear what is actually being designed and implemented it would be really useful to have some clear diagrams that show each service element, its service interfaces and technical bindings, and how it relates to (and uses) other PDS 4 services in the process of delivering user services.

A separate but related issue is that the physical (deployment) architecture was never really shown. I have no clear idea of what class of processors are required, how functions are allocated to them, how failover is handled, what storage is needed, what network bandwidth is required, nor any comparison to what is available. It appears that this is not an issue since we were told that commodity class servers were being used, but fielding a highly reliable distributed system (a design goal) requires that these issues be considered and documented. Similarly there was no discussion of planned system performance.

Recommendation: My recommendations:

- Update the design materials for the architecture to make the actual service architecture and layering clear, derive these from some existing documented method, such as RASDS, SOA-RM, or SCA
- Clarify the actual service interfaces, technical / protocol bindings for each service element
- Document how these elements interact in the delivery of user services
- Define a physical / deployment view for the architecture showing hardware, software deployment, network and storage capacities for all nodes in the distributed system
- Simulate or at least prototype how the distributed system will operate under something like real conditions and evaluate end to end performance

Response:

PDS accepts the recommendations.

Status:

Open

Title:	PDS 2010 Review	RFA:	9	Date:	22-24 Mar 2010
Author:	Peter Shames			Email:	peter.shames@jpl.nasa.gov
Topic:	Improve monitor data (#17 Service Design)				

Comment/Concern: ***Suggest that the PDS should include monitor data reflecting performance of all nodes as a measure of how well they handle requests and a signal of loading impacts and need to upgrade nodes and or links***

Related to the RFA asking that the a priori understanding of the physical / deployment architecture be improved, this RFA suggest that you may wish to evaluate and improve the sorts of monitor data that are to be provided from all of the distributed system physical and service elements. This would include gathering and analyzing a variety of performance information,

Recommendation: My recommendations:

- Analyze the expected deployment and allocation architecture of the PDS 4 system
- Identify resources that need to be monitored
- Identify key performance metrics to be gathered to allow assessment of the health and delivered performance of the system
- Develop any necessary performance requirements (which seem to be missing) and levy them on the system elements as needed
- Develop the necessary monitoring requirements and design the capabilities
- Leverage commercial capabilities as you have elsewhere

Response:

PDS already has plans to capture monitoring and performance requirements when designing PDS 2010 system and science services. It is our intent to leverage commercial capabilities wherever we can.

Status:

Closed

Title:	PDS 2010 Review	RFA:	10	Date:	22-24 March 2010
Author:	Dave Heather			Email:	dheather@rssd.esa.int
Topic:	Management of Information Model (Operation Concept Doc)				

Comment/Concern: Section 5.2.1 of the Operations Concept document outlines the management of the information model. One of the scenarios listed is the 'Delete Operation' in which an existing class can be removed from the model.

I have a big concern here (and likewise with the removal of keywords / permissible values in the dictionary), in that it becomes extremely difficult for data providers to remain compliant if values or classes are completely removed from the model/dictionary. It is especially difficult for long missions where a version of the PDS Standards is selected and 'frozen' at the beginning of the mission. If keywords or classes these data providers use are then removed from the model during the mission, it becomes far more complex to remain compliant and to manage the information model / dictionary versions originally chosen.

No mechanism appears to be in place to manage this sort of situation in the current setup.

Recommendation: Provide a robust versioning for both the dictionary and information model to allow for full validation/ingestion/use of data compliant to older versions in cases where the model evolves during a mission. All PDS4 tools should also be developed to allow for this scenario.

Response:

'Delete operation' is available only for system operations, not information model. PDS accepts the recommendations. Section 5.2.1 of the Operations Concept has been updated accordingly.

Status:

Closed

Title:	PDS 2010 Review	RFA:	11	Date:	22-24 Mar 2010
Author:	Peter Shames			Email:	peter.shames@jpl.nasa.gov
Topic:	Tension between integrated system goals and node autonomy				

Comment/Concern: ***There is a tension apparent between the nodes, who have and want to maintain autonomy, and the PDS users some of whom may want a more integrated system than they have had in the past. Not clear that the right balance has been struck.***

This issue never appeared in any particular set of slides, but emerged during discussions among the board, the PDS EN team, and the DN members. There is no question that the programmatic structure, with the DNs funded separately from the PDS project itself, represents an organizational challenge. It is also seems to be clear that the DNs want to maintain their autonomy, to make their own technical choices, and to meet the needs to their discipline users in the way that seems best to them.

At the same time the PDS 2010 project has goals of:

- ***On-line services allowing users to access and transform data quickly from anywhere in the system***
- ***A highly reliable, scalable computing infrastructure that*** protects the integrity of data, ***links the nodes into an integrated data system***, and provides the best service to both data providers and users

Accomplishing these is going to require a level of integration and standardization, of processes, naming, searching, user interfaces that exceeds what is presently done in PDS 3.

Recommendation: While it is clear that the whole team, EN and DNs has worked together really effectively, it appears that some further levels of integration will be required if these goals are truly to be achieved. My recommendations:

- Review the requirements to make sure that they are sufficient to lend the needed guidance to the architecture and design effort
- Ask the SMD management responsible for the PDS 2010 project and all the nodes to ensure that appropriate guidance is provided to all the elements involved in this project
- Ensure that the integrating technical elements are adequate to the task. This must include: global search terms and language, common GUIs, agreed target names and name disambiguation, integrated search mechanisms, and common data access mechanisms
- Continue the excellent work among all the nodes to ensure that the global goals are met and that local discipline user and provider needs are still satisfied

Response:

We don't preclude changes to that tension. In fact, we design to allow it. PDS 2010 will follow governance processes and system-level requirements directed by the PDS Management Council. These recommendations will be considered by the PDS MC.

Status: Open

Title:	PDS 2010 Review	RFA:	12	Date:	22-24 Mar 2010
Author:	Peter Shames			Email:	peter.shames@jpl.nasa.gov
Topic:	Maintaining consistency among global and specialized schema (#5 Ops Concept, and elsewhere)				

Comment/Concern: ***There is no strict means for managing consistency between user (and provider) derived schemas and new info model, but they expect to have a schema validation capability and may also define a "schema for schemas"***

This topic was discussed twice during the presentations, but the answers still sounded to me like "the hooks are in to support this, but we do not yet have a concrete plan for how to do it". Since there are a lot of opportunities for "XML hackers" to corrupt even the best defined schemas it would be appropriate to design in and agree on some processes for managing this. Similarly, there are lots of opportunities for new names for all sorts of different elements to proliferate. There are mechanisms for managing hierarchical namespaces, but there may need to be a process for identifying and resolving conflicts and promoting key elements to global status.

Recommendation: My recommendations:

- As part of the design develop agreed processes for managing the evolution of the schemas
- As part of the design develop mechanisms for identifying collisions and multiple definitions throughout the namespaces
- Develop processes for resolving ambiguities and collisions and for promoting key terms to global status

Response:

PDS4 model will ensure there's no collision. Generic schema will be used to validate specific missions. Peer review process will be developed to double validate, and resolve ambiguities and collisions. Policies and processes will be documented. Tools will be in place to verify as part of the delivery/ingest process.

Status:

Open

Title:	PDS 2010 Review	RFA:	13	Date:	22-24 Mar 2010
Author:	Peter Shames			Email:	peter.shames@jpl.nasa.gov
Topic:	Support for global name resolver (intro and elsewhere)				

Comment/Concern: ***Not clear that there meta-data and target name consistency that would permit cross node queries to be successful. Not clear how this same problem is going to be handled in the even broader international IPDA case? Will there be some SIMBAD like name resolver (http://simbad.u-strasbg.fr/simbad/sim-fid), which seems particularly useful for all of the small bodies objects?***

This issue also surfaced during discussions among the board, the PDS EN team, and the DN members. Within the astronomy community many different names may be used for the same physical object, for that matter, the same thing happens with spacecraft, where there are MER1 & 2, MER A & B, and Spirit and Opportunity. I suspect the issue may be the most challenging for small bodies where there are a lot of them, they may have different names, and, it appears, some may even change type from comet to asteroid.

Recommendation: My recommendations:

- As part of the design process identify if some sort of global name resolver has a place in the architecture
- Determine if there needs to be one for the whole PDS 2010 or ones for specific DNs
- Define a global mechanism that makes this service available, ala Simbad, for use in queries
- Do the work to populate it (them)

Response:

PDS 2010 Project defines a core gold source in its data dictionary for names (e.g. Missions). It also allows for synonyms /aliases to resolve other names used outside of PDS. This RFA raises two separate issues, 1) how are alternate names captured and 2) how are the names resolved?

The PDS4 data standards will address the first issue in two ways. The information about a target body, mission, instrument or any other entity, including its name(s), will be captured in a PDS4 product and registered with the PDS 2010 registry. Within a product, the name(s) are captured in the attributes “title” and “alternate title”, where “title” is the currently preferred name, and “alternate_title” is a list of other names. Since any PDS4 product can be versioned, a new name, possibly the preferred, can be added. In addition, the data dictionary allows multiple names. For each name the dictionary also provides the “begin and end” effective dates, natural language used, and meaning.

Name resolution will be handled in a number of ways. The PDS 2010 architecture includes a metadata harvester. For any entity, for example target bodies, the harvester can extract names, from the title, alternate title, and data dictionary (i.e. obsolete names) and provide them to search and name resolution services. For example a text/facet based search application (e.g. solr) would support the search on any extracted name and return a link to the associated PDS4 target product. The PDS Small Bodies Node has developed the target name resolver service. This service is able to resolve the more complex naming schemes used for comets and asteroids, for example where a particular body changed classification from comet to asteroid and had several names over time.

The design, deployment, and interaction with other existing name resolvers (e.g. Sinbad) will be addressed for Build 2 of PDS 2010.

Status:

Closed

Title:	PDS 2010 Review	RFA:	14	Date:	22-24 March 2010
Author:	Dave Heather			Email:	dheather@rssd.esa.int
Topic:	Validation layer for Node-level requirements in the standards				

Comment/Concern: There was a concern raised during the review board discussions that there is very little structure/architecture provided for validation of the changes made to the Dictionary and Information Model by the individual Nodes.

One of the issues with PDS3 has been that the Nodes have interpreted the Standards differently from one another, and each has entered their own keywords to the Dictionary or placed restrictions on values / keywords used. Over the years this has produced a system in which data providers are given different requirements for the same data product depending on the Node to which they are delivering. Similarly, data users are presented with products that can vary in content quite significantly depending on the Node (or international partner) from which data are retrieved.

The PDS2010 system presented seemed to provide a good deal of flexibility in the dictionary and information model for use by the Nodes, which is good. However, no mechanisms seemed to be in place to validate these changes and avoid running into the situation we now have in PDS3. It is possible that, to a certain extent, this can be avoided by having a very robust PDS4 Standards Reference document, but this was not available at the time of the review. It is also not clear how the Standards Reference will provide both a clear list of requirements for PDS compliance and the specialized (and evolving) requirements of a Node. It is therefore also worth considering the possibility to add a more automated validation layer at the top-level for the changes requested by Nodes, ensuring core PDS-compliance requirements are not compromised.

Recommendation: Three recommendations arise from this concern

- Ensure that this is considered as the design process continues
- Provide a very robust Standards Reference document that will ensure that the specialized requirements of the individual nodes can be met without compromising the core requirements.
- If at all possible, put in place an automated validation layer that will ensure changes made by Nodes will not compromise the core standards

Response:

Recommendation 1

The PDS4 data standards address the ambiguity and variability in the meaning and use of data elements (aka keywords) primarily through the use of a single shared model. Even though a node or a mission can propose a new data element and manage it at the level of a node or mission, the new data element must be consistent within the single PDS information model. For example request for a new “latitude” data element will be evaluated with respect to its relationship to the existing “latitude” data element. In the majority of cases, it is expected that the request is actually for a variation of an existing more general data element and

a formal “extension” or “restriction” can be performed. A data element that is unique will simply be created with a unique name and defined accordingly. A data element classification scheme is also being developed to help data providers identify existing keywords that can be re-used with or without extension.

The PDS4 data standards will also allow the adoption of data elements from other models external to the PDS. Duplicate data elements might exist however a separate namespace grouping will indicate that they are entirely different from any PDS data element.

Recommendation 2

The Standards Reference document is being written to be in conformance with the PDS information model. Standards that can be modeled will first be defined in the PDS information model. Subsequently the standards reference document will provide additional policy and guideline information for their use. Standards that cannot be defined in the PDS information model will be documented in the Standards Reference document.

Recommendation 3

To ensure that the core standards are not being compromised, the information model is being used to generate the core XML schema. The core XML schema are then specialized by the nodes and mission for mission products. The core XML schema can be used to validate the mission products to ensure that the core standards are not being compromised. A product conforming to a specific schema must also conform to the generic schema from which the specific schema was derived.

Status:

Closed

Title:	PDS 2010 Review	RFA:	15	Date:	22-24 Mar 2010
Author:	Peter Shames			Email:	peter.shames@jpl.nasa.gov
Topic:	Primacy of Information Model (#9 PDS Data Architecture)				

Comment/Concern: ***#9, pg 7, there is an issue in that the info model is actually defined in an ontology, but the slide says "implemented as XML schema", this should be made clear***

From discussion it became clear that the authoritative information model for the data architecture is actually an ontology maintained using an open source tool, Protégé. This slide obscures that fact and it is really never made clear in these materials even though it is alluded to. This approach, of using a formal ontology and deriving schema, documents, and data structures from it is a real strength of the design and it should be acknowledged and taken advantage of wherever it can be.

This approach can be somewhat daunting to those unfamiliar with it, but it is very powerful and the team is to be applauded for using it and for finding ways of making derived materials accessible.

Recommendation: My recommendations:

- Update the design materials for the architecture to make this approach clear
- Leverage the approach wherever it makes sense to do so, since it is a real strength for maintaining conceptual and expressive clarity
- Sustain the processes for translating these core design materials into more easily understood and accessible forms for others to use

Response:

PDS agrees with the recommendations and is already addressing them.

Status:

Closed

Title:	PDS 2010 Review	RFA:	16	Date:	22-24 March 2010
Author:	Dave Heather			Email:	dheather@rssd.esa.int
Topic:	Tools planned for PDS4				

Comment/Concern: Tools presented during the review did not include the following, which should be considered in order to provide the best service to data providers and end-users:

- Conversion tools from PDS3 to PDS4 *and* from PDS4 to PDS3
- Porting of PDS4 data to the most commonly used scientific packages. Providing a replacement for NASAView style visualization is good, but scientists would want codes that would allow for software such as IDL, Envi and ISIS to open PDS4 data.

During the review, Slide 29 of the System Architecture presentation listed a large number of 'Science Related Services' including:

- Coordinate System Transformation
- Calibration on the fly
- Map overlays
- All-Purpose geometry engine

I saw no associated plans for this sort of service in the documentation and the nature of these services would make it extremely difficult to incorporate into a general PDS tool-suite.

Recommendation: Consider provision of tools for PDS3<->PDS4 transformation and for porting to the most used scientific packages.

The science related services seem to me to be primarily node-related, and should probably not be considered as part of the top-level PDS tool suite.

Response:

The information presented during the System Review was a high-level summary of science-related services and product transformations that had been identified previously by PDS Discipline Node staff. The detailed list of transformations includes PDS3 to PDS4 and back as well as transformation to formats supported by scientific packages. The detailed list includes prioritization and will be implemented per the schedule and resource availability.

Yes, many of the science-related services will be Node-specific. At this stage in the design, they have simply been identified and provided a place within the architecture where they will reside.

Status:

Closed

Title:	PDS 2010 Review	RFA:	17	Date:	22-24 March 2010
Author:	Dave Heather			Email:	dheather@rssd.esa.int
Topic:	Tool Distribution - comment				

Comment/Concern: The reasoning behind distribution of tools to Nodes is unclear in many instances.

Registry Service SRD/SDD Section 6.1: It is not clear to me why the preferred deployment for the Registry Tool should be locally at Node level. If the objectives for PDS2010 include providing access to any PDS data from anywhere within the system, and moving the nodes into an integrated data system, it may be more logical to have a single Registry for the entire PDS holdings, automatically updated each time the Harvest Tool is run at each Node. I see no advantage to running the tool at the Nodes above having a centralized instance that will aid in the objectives noted above.

Report Service SRD/SDD Section 8: Why is there only a single EN instance of the Report Service? The paragraph says '…This will allow each Node to submit their metrics to a centralized location.' Does this mean the Nodes will run the tool remotely from the EN to obtain and submit their metrics?

Recommendation: Clarify the reasoning behind these choices in the documentation and/or re-think the distribution policy as needed.

Response:

The history of the PDS is one of a distributed system and it is not the intent of PDS 2010 to centralize the system but to bring federation to the distributed Nodes. By having a local instance of the Registry Service at each Node, we foster ownership and stewardship of the information they place in the registry. In addition, several of the Nodes have local catalog systems that will live on past the deployment of PDS 2010 and having the Registry Service installed and running locally increases options for integration with those systems. We can still meet the stated objectives through population of the aggregate instance of the Registry Service as specified in the design.

Unlike the Registry Service, the Report Service will most likely be satisfied with a commercial software package. Early on in the architecture development phase we assumed that licensing costs and possible issues with platform support made a centralized service solution more desirable.

Status:

Closed

Title:	PDS 2010 Review	RFA:	18	Date:	22-24 March 2010
Author:	Dave Heather			Email:	dheather@rssd.esa.int
Topic:	Standardization of data access / web pages at nodes				

Comment/Concern: It was mentioned during the presentations that there are significant efforts underway to improve the look and feel of websites as part of the phasing in of PDS 2010.

It would benefit end-users and newcomers to the PDS if the look and feel of the web pages at each Node, especially the data access pages, could be standardized as much as possible. It is understood that all Nodes will have their own requirements to streamline searches against their own specialized data holdings. However, one of the aims presented during the review was to provide

- ***On-line services allowing users to access and transform data quickly from anywhere in the system***
- ***A highly reliable, scalable computing infrastructure that*** protects the integrity of data, ***links the nodes into an integrated data system***, and provides the best service to both data providers and users

A standard look and feel across all PDS sites would be very important in this case.

Recommendation: As much as possible without compromising the needs of the individual Nodes, standardize the look and feel of the websites across the PDS.

Response:

The recommendation will be considered by the PDS MC.

Status:

Open

Title:	PDS System Review	RFA:	19	Date:	22-24 Mar 2010
Author:	Dave Linick			Email:	Terry.D.Linick@jpl.nasa.gov
Topic:	Requirements Specification				

Comment/Concern: The flexibility inherent in this architecture provides hooks for its future evolution and recognizes that not all of the customer needs/desires can be accommodated in the timeframe of PDS-4 development (October, 2011). However, it is not always clear what specific capabilities will be provided on that time frame, and what will be added later as time and resources permit. This can lead to misalignment of expectations.

1. **Recommendation:** I recommend that the specific requirements that will be satisfied by October, 2011, be crisply defined by the project and agreed to by the sponsor. This provides a measurable way of assessing the success of the project and assures that the project is clear about what is expected.

   Perhaps one approach to this that recognizes the uncertainty in the required implementation effort and schedule is to prioritize the requirements and indicate which of those the project commits to have completed in the required time frame. The remaining requirements can be satisfied if time and resources allow, but are not "guaranteed." They also provide a blueprint for future evolution.

Response:

Level 3 requirements traceability to PDS 2010 Project can be found at: http://pds-engineering.jpl.nasa.gov/meeting/systemreview/UsabilityOfRequirementsInPDS3_and_PDS4.pdf. It shows what are the specific requirements will be satisfied by Oct 2011, and the remaining requirements will be prioritized and planned.

Status:

Open

Title:	PDS System Review	RFA:	20	Date:	22-24 Mar 2010
Author:	Dave Linick			Email:	Terry.D.Linick@jpl.nasa.gov
Topic:	Transition Requirement				

Comment/Concern:

It is unclear when project and data providers must use the PDS-4 format.

2. **Recommendation:**

   Define and communicate a crisp statement of the requirement to comply with the PDS-4 format.

Response:

The recommendation is under consideration by the MC.

Status:

Open

Attachment 2

RFA Summary

Total RFAs : 20

Total Open pending for resolution / completion : 6

Total Closed : 14

RFA #	Topic	Author	Status	Comments
1	The core data model and data dictionary are crucial for distributed queries	J. Merka	Closed	Recommendations are already in work or in plan
2	Data node holdings overlap in content	J. Merka	Closed	Clarification provided
3	How much NASA funding goes towards supporting collaboration with international partners?	J. Merka	Closed	IPDA support resource allocated by nodes
4	What is PDS2010 relationship to the NASA Virtual Observatories?	J. Merka	Closed	VxOs relationship clarified
5	Documentation Inconsistencies and comments	D. Heather	Closed	Ckarified and documents are updated
6	Support for global data searches (#1 intro, #16 architecture, and elsewhere)	P. Shames	Closed	Search Service planned in Build II
7	Harvest and Registry Tool Comments	D. Heather	Closed	Clarification provided
8	Improve architecture description and understanding (#17 Service Design)	P. Shames	Open	Recommendations accepted
9	Improve monitor data (#17 Service Design)	P. Shames	Closed	Recommendations in plan
10	Management of Information Model (Operation Concept Doc)	D. Heather	Closed	ops concept updated
11	Tension between integrated system goals and node autonomy	P. Shames	Open	MC to consider recommendations
12	Maintaining consistency among global and specialized schema (#5 Ops Concept, and elsewhere)	P. Shames	Open	Close when processes are defined
13	Support for global name resolver (intro and elsewhere)	P. Shames	Closed	Clarification provided

14	Validation layer for Node-level requirements in the standards	D. Heather	Closed	Clarification provided
15	Primacy of Information Model (#9 PDS Data Architecture)	P. Shames	Closed	Recommendations are already in work or in plan
16	Tools planned for PDS4	D. Heather	Closed	Recommendations in plan
17	Tool Distribution - comment	D. Heather	Closed	Clarification provided
18	Standardization of data access / web pages at nodes	D. Heather	Open	MC to consider recommendations
19	Requirements Specification	D. Linick	Open	Close when requirement traceability matrix in place
20	Transition Requirement	D. Linick	Open	MC to consider recommendations

Attachment 3

Responses to questions and recommendations provided by board member, Dave Korsemeyer

1. **Q: Is the prime goal long term storage or access to data?**

The prime goal is to preserve the data and provide online access to them to support NASA's R&A efforts.

2. **Architecture and design were presented as "in sync" and virtually synonymous. I don't know which one drove the other. This was not clear.**

Design is driven by the Architecture. PDS 2010 Project had an explicit architecture phase that preceded the start of the design phase.

3. **How does the PDS4 design compare and contrast with the LMMP work happening in ESMD?**

These two projects have strong collaboration and leveraging since some of the PDS 2010 Project team members are concurrently working on the LMMP. Wherever possible, design and applicable capabilities are shared.

4. **We did not see alternative designs discussed or what downselect process occurred. That would have been enlightening to assess whether the design was as robust as possible.**

PDS 2010 Project holds on-going design working group meetings and periodic internal technical sessions involving all PDS nodes to discuss alternative design choices and trade-offs. Those discussions and decisions are documented on the working group wikis and technical session notes. The project will be happy to provide those documents.

5. **Little justification on resources. Was the work scaled to fit the budget or the budget scaled to fit the work?**

PDS defined its vision, drivers and requirements first. Then, the PDS 2010 Project plan was constructed based on negotiation with NASA HQ and the PDS MC included available resources (funding and staff). PDS 2010 Project is funded by the NASA Headquarters as "over-guide" funds through the PDS task funding vehicles to the PDS Nodes as proposed in the POP. The resource allocation is based the "in-guide" repurposed to support the PDS 2010 project by the way of reprioritizing the key staff and tasks.

6. **Who were the 3rd party reviewers of this plan other than this board?**

Other 3rd party reviewers include the PDS Management Council, other PDS staff outside of the PDS 2010 Project working groups, and other JPL technical staff.

7. **Are each of the subprojects going to have Preliminary Design Reviews, CDRs, and then delivery reviews? Would like to see some more rigor in the PDS Project management with respect to the technical review process.**

An appropriately tailored NPR-7120.8 is being applied to PDS 2010 Project including review process. A mapping of NPR-7128.8 requirements and PDS 2010 Project Plan is included (see Attachment 1).

Attachment 4

Response to questions and recommendations provided by Board Member, Andy Schain

1. **Is there a formal specification for the data dictionary and a process for keeping it updated?**

Yes, there is a formal specification for the data dictionary. The process for keeping it updated is being worked and will be documented as one of the Project's deliverables. Currently, in PDS3, there is a formal process.

2. **Recommendation: Consider having the data dictionary contain terms from previous versions of PDS. Consider a mechanism to computationally resolve semantic conflicts that incrementally becomes a richer service.**

First, the data dictionary will contain terms from previous versions of PDS. Secondly, we maintain a valid ontology and resolve semantic conflicts. We will work on an effective mechanism.

3. **Will the dictionary output be available for other systems (e.g. LMMP, Moon in Google Earth) to use as part of a query or search service? Recommendation: If not already being planned, consider a WSDL service for downstream customers.**

The dictionary is extracted in RDF. Online web services available for other systems can be considered after PDS 2010 system is operational (after Build 2).

4. **Recommendation: Target specific delivery and response goals in assisting those who must generate XML schemas as part of their PDS submission. Consider mechanisms to identify differences in logic or attributes of classes in XML schemas and process to resolve them. Recommendation: Plan this activity in your project plan.**

There are many ways to interpret XML schema, we are working PDS-wide approach and best practice for use of XML.

5. **There is a strategy for expressivity and RDF seems to be the choice for the information model and the data dictionary, yet the documentation seems to be fairly shy about that. Recommendation: Be specific.**

PDS 2010 Project accepts this recommendation as specified.

6. **Recommendation: Develop a concept of operations inclusive of certain types of change requests (e.g. conflict resolution) within your project plan.**

PDS 2010 Project accepts this recommendation as specified.

7. **Recommendation: Consider adopting as a starting point the provenance attributes identified in the NASA Information Architecture Policy Guide.**

PDS 2010 Project accepts this recommendation as specified.

8. **Recommendation: Any additions or holes in the NIA should be identified and sent to the author prior to the end of FY 2010.**

PDS 2010 Project accepts this recommendation as specified, and will consider it.

9. **Will there be a common query engine or a common query service that nodes can use to build additional query browsers or interfaces? Recommendation: Specify the architecture and make it part of a focused review.**

A common query engine and query service will be available for the nodes to build additional query browsers or interfaces.

10. **Recommendation: Perform the analysis to determine advantages, disadvantages, trades for a uniform storage solution that provisions analytic cluster/cloud services (e.g. Hadoop), and provisions for archival needs.**

The PDS Physical Media Working Group has begun an effort to perform this type of analysis.

Attachment 5

NPR 7120.8 / PDS 2010 Project Mapping

	NPR 7120.8 requirements	PDS 2010 Project Plan	Description
4.5.1.1	Project Lead shall establish a WBS	Section 3.3.1	PDS 2010 Project WBS includes Management, Systems Engineering, Data Standards , System Development and Operations
	Project schedule with milestones for each element in the WBS	Section 14.2	PDS 2010 Project's schedule with milestones are posted on http://pds-engineering.jpl.nasa.gov/index.cfm?pid=100&cid=118
	An allocation of the project's available resources necessary to achieve each milestone	Section 14.1.1	The PDS 2010 project is funded by the NASA Headquarters funds through the PDS task funding vehicles to the PDS Nodes as proposed in the POP. The resource allocation is documented in the POP.
	The milestones should be chosen at intervals sufficient to demonstrate steady progress towards achieving the overall KPPs for the project	Section 12.5 - 12.8	The PDS 2010 project progresses in releases. Release phasing is summarized in Project Plan section 12.5.. Section 12.6 - 12.8 describes the details of each release.
4.5.1.3	Project Lead shall track progress against a baseline plan. The WBS, the project schedule, and the allocation of resources to milestones constitute the baseline plan for assessing technical, schedule, and cost performance.	Section 6	A monthly report is provided to the PDS Program and NASA Headquarters Management. The progress and risks of the project are presented to the PDS MC on the monthly MC teleconference and at each MC face-to-face meeting. In addition, status is reported to the PDS Program Manager and Chief Scientist on regular monthly Engineering Node teleconferences.
4.5.2.1	Program Lead may authorize special independent assessments at any time in a TD Project's life cycle.	Section 14.2.4	Key reviews have been identified as special independent assessments. Schedule of key reviews is listed in Project Plan section 14.2.4.
4.5.2.2	The Project DA shall determine if the optional KDP (KDP B) is required during Formulation or if the optional KDP (KDP B) is not needed	Section 4	During the project definition phase, several documents including the PDS 2010 level 1, 2, and 3 requirements, and white papers were produced to support Key Decision Points.
4.5.2.3	The Project DA shall determine if optional KDPs (KDP D and E) are required during Implementation or if the optional KDPs (KDP D and E) are not needed.	Section 6	A monthly report is provided to the PDS Program and NASA Headquarters Management. The progress and risks of the project is presented to the PDS MC on the monthly MC teleconference and at each MC face-to-face meeting. KDPs will be determined by the MC during Implementation phase.

4.5.2.4	Independent Assessments (IAs) occur as part of the TD Project life cycle. IAs during Implementation are performed periodically and should be documented in the Project Plan.	Section 14.2.4	IAs have been identified as special independent assessments during implementation. Schedule of IAs is listed in Project Plan section 14.2.4.
4.5.3.1	The TD Project Lead shall conduct TD Project status reviews annually to assess both progress towards the KPPs and the maturity of the technology.	Section 6	A monthly report is provided to the PDS Program and NASA Headquarters Management. The progress and risks of the project are presented to the PDS MC on the monthly MC teleconference and at each MC face-to-face meeting. In addition, status is reported to the PDS Program Manager and Chief Scientist on regular monthly Engineering Node teleconferences.