February 12, 2008

PDS4 Discussion Meeting PPI/EN Notes

Attendees: Dan Crichton, Steve Hughes, Steve Joy, Todd King, Ray Walker

The meeting reviewed a concept outline (PDS4 Concept Outline, 2/7/2008) that the PDS4 Planning WG has been putting together. The following is notes regarding the discussion of several of the items in both the outline and the PDS4 schedule. In addition, some missing items were identified.

1. Vision – Regarding the vision, there was extensive discussion that PDS4 should represent a paradigm shift to offer both new and better "services" for users. Users should be able to find, access and retrieve data in formats which are useful to them. Currently, grant funded scientists have such constrained budgets that spending time converting data to useful formats is a problem. It was also discussed that PDS does have a "preservation" responsibility, but it needs to step up to provide support to users who are accessing PDS to do science.

2. Unambiguous Standards – There was general agreement that this is the number of priority of PDS4. In addition, Ray felt that PDS should try and identify a core set of formats such as image, table and cube. It was discussed the PDS should be "forward compatible" or can be "re-configured" to work with PDS4, but not "backward compatible" meaning that we should be able to take PDS3 data and covert to PDS4.

3. International Archives – We should be explicit that a PDS4 goal is ensuring that we can link with international archives and that discipline-specific users can access science data for a discipline across international boundaries.

4. Distributed Online System – PDS should be built with a strong foundation. The distributed infrastructure needs to be a core set of functions on which higher order PDS-specific archive and user services are built. In addition, search and access should be to the product, record and/or catalog level that may differ across discipline nodes.

5. Integration with Data Providers – There was general discussion that PDS needs an OLAF-type capability for DAP-type proposals.

6. Automation – It was discussed that the desired stated of PDS4 is one that supports both "efficiency" in operations and "agility" in terms of processes.

7. High Speed Data Exchange – It was discussed that this should cover all aspects of data delivery including both context and transfer as well as interfaces to/from PDS (i.e. missions to PDS, PDS to users, PDS to NSSDC)

8. Tracking/Reporting – It was discussed that PDS should engineering a "reporting" capability into PDS that derives reports on status of data and metrics on usage.

9. Flexible Search – Ray brought up the question of how do we really meet the Arvidson vision of "one-stop" shopping. It was discussed that search needs to work with the distributed online infrastructure to ensure that data across nodes can be located, accessed, and then displayed in "discipline-specific views". These discipline-specific views allow data from multiple nodes to be presented in a form that is useful to that DN user.

10. Coordination and Management: It was discussed that PDS needs to be efficient, as was mentioned earlier, but also needs to be well managed with clear delegation of authority. This was also discussed in the context of the PDS information model where a possible re-organization of the model with "namespaces" might help in partitioning it so that disciplines could manage parts of the model that is unique to them.

11. Highly Scalable, Reliable Computing Infrastructure – It was discussed that PDS needs to have a reliable infrastructure that users can depend on. PDS4 should make investments in hardware and operations that help us achieve a highly available up time. PDS also needs to have a plan that allows it to scale as the usage and data scales.

12. Knowledge-base: This was discussed and determined that the description should be re-written from what currently exists. It was discussed that having a "growing" knowledge-base that users get to and use data would be helpful as PDS improves its services.

13. Schedule – Regarding the schedule, there was discussion regarding the transition and how to go from PDS3 to PDS4. Ray felt that it was important that we not "bandaid" PDS3 by trying to fix/port legacy software. We should implement PDS4 making good design decisions based on the design principles that PDS has been discussing. We should identify when we cutover in the schedule. We should maintain the PDS3 system until we are ready for that cutover. In addition, Steve Joy requested that we show more detail on the schedule for the projects. There was also discussion regarding the need to organize the schedule so that the projects appropriately build upon one another.

14. Services – There currently is no discussion on some of the core services. It was recommended that some core services be defined (e.g., format conversion, coordinate translation, stretch, subset, superset, etc)