

PDS4 Core January 2013

Comments from Simpson

The document is surprisingly 'soft' and incomplete. Having worked with Steve, I had expected more detail on the Information Model; the list of software, tools, and services is very qualitative and many of the "descriptions" are actually just restatements of status.

1. The first two sections ("Introduction" and "Core Information Model") are not helpful and could be eliminated. For example, MC does not need definitions of base data structures; and the list of "subclasses" is (at best) informal and not obviously related to what will be defined later as the Core (*e.g.*, PDF/A in the first table is a coding scheme, not a subclass). Specification of the Core classes in Appendix A is a better place to start.

2. MC needs to add an explicit statement at the beginning of the document. Is the Core intended to support LADEE and MAVEN (only), is it intended to support data migration, is it intended to support any other missions (*e.g.*, OSIRIS REx) and/or R&A activities, or is it some or all of these? The Information Model Core appears to be predicated on the LADEE/MAVEN assumption (Section II, paragraph 1); the Label Generation software will support migration only. A clear statement of why we need the Core will help in determining what is in it and then the implementation priorities.

3. The Core class specification is not detailed enough. Field_Bit, which has been repeatedly requested by PPI and is two levels below Field_Binary, is not shown in Appendix A. Steve has one spreadsheet which lists exactly all of the classes which are in the Core; he has a second which lists all of the classes in the Information Model and indicates whether each is in or out of the Core. At least one of these should be appended to the document. Without a complete list of included and/or excluded classes, there will be disagreement down the road about where effort should be placed. Several of the lower level classes are "optional" (including Field_Bit) and these could be particularly contentious.

4. There are errors in Appendix A. The legend shows a category "Requires Advocate" (white); there is no class in the following table in this category. "Candidate V1.1" has a projected release of "Build 4a" where a date would be more appropriate. More importantly, Product_SPICE_Kernel is needed if we are to have SPICE kernels in any PDS4 archive; it is not shown in Appendix A although I pointed this error out previously and was told it would be fixed. In fact, I wonder whether this version of the document is the final or whether some intermediate version was accidentally distributed. For what it's worth, I have checked Steve's supplementary (inclusive) spreadsheet and it appeared to be OK; so the errors appearing in Appendix A are not in the underlying data base.

5. Whether "Candidate V1.0 – For Internal Use" is part of the Core is ambiguous. Although these classes — and the data products and metadata based on them — are firmly under centralized control, it would be helpful to know whether they are considered to be part of the Core; the release date is the same.

6. Section III begins with “The software being proposed ...” is not a very strong commitment. Are these pieces of software, tools, and services required to meet the objectives stated in the preamble (after the upgrade suggested in #2 above)? If so, maybe they need to be more than “proposed.”

7. If this software is required, what does each piece do? “Provides functionality” is not a strong statement. It is also not clear whether the listed tools and interfaces are part of the services listed at the top of each software table or whether tools, interfaces, and services are all parallel. For example, does the Registry Service provide the Data Registration function, and its major components are the Harvest and Catalog Tools; or is Registry Service complemented (somehow) by Harvest and Catalog?

8. Are Label Generation, LDD Generation, Label and Data Validation, Label and Data Transformation, and the Bundle Browser part of the Core? I thought these were important functions, which have been endorsed at both the DDWG and MC level; yet the statement toward the bottom of page 3 implies otherwise: “In addition to the core system software, the following functions provide support for preparation, review, and use of PDS4 data.”

9. What is a “proxy product” (Data Registration) and a “facet-based search” (Data Search)?

10. Why are two software developments supporting migration when there is none supporting new PDS4 label generation?

11. The “description” of LDDTool and Validate Tool is a status; if these pieces of software are required for the Core, is there not a set of requirements listing what each should do? A summary of those requirements should be listed under “Description”. A similar comment applies to other software “descriptions,” most of which could be beefed up.

12. The Format Transformation sheet (Label and Data Transformation) has not been included with the document. Even if it were, what we need is the list of transformations *required* for Release 1.0, not the list of requested transformations.

13. The PDS4 equivalent of the PDS3 OAL does not obviously belong under Label and Data Transformations. I would classify it as a ‘toolbox’ that probably supports many of the other functions (a separate category). Some of the toolbox inhabitants may be more important for Release 1.0 than others; have these been prioritized?

14. The Bundle Browser doesn’t appear to have components yet; it is not clear what it should do. I am having a hard time imagining that something so ill-defined in January will be ready for testing in March and deployment in July given all of the other work to be done.

15. The list of documents is helpful; but is there any information on how the various pieces fit together? I’m working on this myself and don’t yet have a good picture of the final relationships.