PDS Tool Working Group (PDS-TWG)

PDS MC F2F

April 20, 2017

Agenda

- Charter, Members and Logistics
- Level 3 Requirement Addition
- State of Tools
- Tool Registry
- Transformation Prioritization
- Validation Prioritization
- Next Steps

Charter

- Charter
  - Maintaining the PDS Level 3 requirements for tools;
  - Reviewing and overseeing the PDS-wide tool inventory;
  - Identifying gaps and overlaps in tools;
  - Recommending tool priority and phasing for development;
  - Reviewing and commenting on specific tool requirements;
  - Maintaining PDS-wide Tool Schedule and Plan;
  - Supporting beta testing of tools as part of a release cycle;
  - Hosting a tool summit with the PDS technical group; and
  - Provide regular reports at the Management Council Face-to-Face meetings.
- Will be renewed annually

Members

- D. Crichton, Chair (E. Law as alternate)
- E. Guinness, Geosciences
- Joni Johnson, Atmospheres
- Todd King, PPI
- Jordan Padams, IMG (Moses Milazzo as alternate)
- Eric Palmer, PSI
- Tanya Lim, PSA

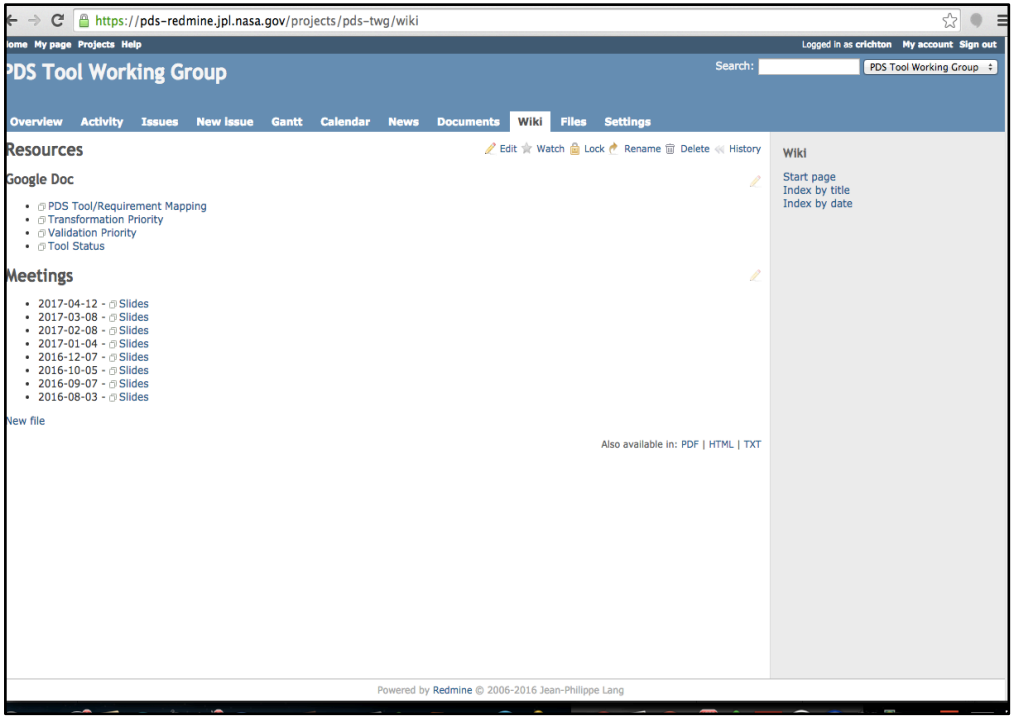
* Sean Hardman, EN as ex-officio advisor

Logistics

- Mailing list: pdstwg@list.jpl.nasa.gov
- Monthly telecon via webex started in Aug 2016
  - 1st Wednesday 7:30 am Pacific Time
  - Eight meetings held to date
- Work spaces:
  - https://pds-redmine.jpl.nasa.gov/projects/pds-twg
  - https://docs.google.com/spreadsheets/d/18oqtg3DEo2KrgvBOWLSOuqF2uZtq2XmByJwUknYSZUQ/edit#gid=126564646

Introduction

- The working group has focused on three activities to support current and future builds
  - Identifying the list of tools being developed within PDS and capturing capability and support plans
  - Identifying and classifying requirements for additional validation (including build dates)
  - Identifying and classifying additional transformations beyond the current set
- The working group also mapped level 3 requirements to tools



PDS Tool WG Wiki

Level 3 Requirement Addition

- Based on the requirements to tool mapping exercise, the team determined a gap with respect product design.
- The following requirement was added to the PDS requirements document:
  - 1.5.5 - PDS will provide tools to assist data producers in designing PDS compliant product labels.
  - Would like to discuss approval of requirement change in Exec Session.

PDS Tool Activities

- The team is tracking status for actively developed tools at the EN and the DNs.
- A traceability matrix was set up to map Level 3 requirements to Tools
- Nodes were asked to identify tools beyond the list (on-going request)
- This was facilitated with the following Google doc:
  - https://docs.google.com/spreadsheets/d/18oqtg3DEo2KrgvBOWLSOuqF2uZtq2XmByJwUknYSZUQ/edit#gid=1340660799

Tool Name	PDS Version	Developer/Maintainer	Current Status	Future Plan
PDS Label Assistant for Interactive Design (PLAID) (Formerly APPS LDT)	PDS4	AMMOS	Released	New features planned and in development.
Educational Labeling System for Atmospheres (ELSA)	PDS4	Atmospheres	In Development	
Generate Tool	PDS4	Engineering	Released	Routine maintenance. Plan to merge with igpp.docgen.
igpp.docgen	PDS4	PPI	Released	Routine maintenance. Plan to merge with Generate Tool.
MakeLabels	PDS4	Geosciences	In Development	
LDDTool	PDS4	Engineering	Released	Updated with each build.
Validate Tool	PDS3 PDS4	Engineering	Released. Subtantial support for syntatic and semantic validation.	Focus on content validation. Updated with each build based on TWG prioritization.
Validation Tool (VTool)	PDS3	Engineering	Released	Replaced by Validate Tool.
PDS3 Volume Validator	PDS3	ARC/Engineering	Released	Decommission due to IT security issues.
Transform Tool	PDS4	Engineering	Released. Support for a core set of transformations.	Focus on expanding beyond core set. Updated with each build based on TWG prioritization.
PDS4 Tools	PDS4	Engineering	Released	Routine maintenance.
pds.cdf	PDS4	PPI	Released	Routine maintenance.
ReadPDS4 for IDL	PDS4	SBN	Released	Routine maintenance.
ReadPDS3 for IDL	PDS3	SBN	Released	Routine maintenance.
InspectTool	PDS3 PDS4	Engineering	Requirements/Design	Design underway. Initial release planned for Build 8a at end of FY17.
PDS4 Viewer	PDS4	SBN	Released	Routine maintenance.
NASAView	PDS3	Engineering	Released	Routine maintenance. Plan to replace with Inspect Tool.
PDS3 Table Slicer	PDS3	ARC/Engineering	Released	
Integrated Software for Imagers and Spectrometers (ISIS)	PDS3	Imaging/USGS	Released	
Online Archive Facility (OLAF)	PDS3 PDS4	SBN-PSI	Released	
AMMOS-PDS Pipeline Service (APPS)	PDS3 PDS4	AMMOS	Released	

Tool Registry

- The application is available at:
  - http://pds.nasa.gov/tools/tool-registry
- Migrations of the PDS Search Tools and IPDA Tools have been completed.
- Beyond migrated entries, approximately 25+ new entries have been captured (current total of 69).
  - No new entries in the last couple of months.
- Submissions have been reviewed by the TWG.

Transformation Prioritization

- Today, the Transform Tools supports approximately 27 transformations as of the Build 7b release.
- The team was tasked with adding to and prioritizing the list of supported transformations for the Transport Tool.
- This was facilitated with the following Google doc:
  - https://docs.google.com/spreadsheets/d/18oqtg3DEo2KrgvBOWLSOuqF2uZtq2XmByJwUknYSZUQ/edit#gid=635102082

Input	Output	Comment	Status	Priority (High, Medium, Low)							Score
				ATMOS	GEO	IMG	PPI	RMS	SBN	PSA	
PDS3 Image (8-Bit Image) (16-Bit Image)	GIF, JPEG, JPEG 2000, PNG, PNM, TIFF	[EN] These output formats are included in the Java Advanced Imaging library which is included in the MIPL Transcoder library.	Included in Build 7b								
	PDS4 Labeled Image	[EN] Need to break this out into specific PDS4 product types. [IMG] Priority for GDAL is a PDS4 reader not currently a writer. There are plans for an ISIS3 writer.	Planned for Build 9b	Medium	Medium	Medium	Low(*)	High	Medium	Low	11
PDS3 Table	CSV	[GEO] Would like to see a focus on binary table conversion. [PPI] Already have tools to perform this transformation (for both ASCII and binary).	Planned for Build 8a	High	High	Medium	High	Medium	High	Medium	15
	PDS4 Labeled Table	[EN] Need to break this out into specific PDS4 product types. [IMG] GDAL only has a PDS3 table reader. Can write about 20+ "table" formats (e.g. csv). If Latitude and Longitude are headings in PDS3 table, it will support writing out geometry/GIS point formats (GML, shapefiles, KML,… http://www.gdal.org/drv_pds.html) [PPI] PPI has an existing tool (igpp.docgen) to transform a PDS3 label to a PDS4 label, but external information is needed to create a label that takes advantage of the full capabilities of PDS4.	Planned for Build 8b	High	Medium	Medium	High	Medium	High	Low	13
PDS4 Array* (Array_2D) (Array_3D)	FITS	[EN] This should work on anything derived from an Array_Base. See Binary Data Type Conversion below for details.	Planned for Build 10a and beyond	High	Low ***	Low	Low(*)	High	Low	Low	10
PDS4 Binary Table* (Table_Binary)	CSV		Included in Build 7b								
	FITS	[EN] See Binary Data Type Conversion below for details. [PPI] PPI sees negative value in converting tables to FITS format.	Planned for Build 10a and beyond	High	Low ***	Low	Low	Low	Low	Low	8
	PDS4 Character Table	[PPI] CSV meets the need for binary to ASCII conversion.	Planned for Build 8b	High	High	Medium	Low	Medium	Low	Medium	13
PDS4 Character Table** (Table_Character)	CSV	[EN] PDS Table Character to selected standard implementation(s) of CSV.	Included in Build 7b								
	FITS	[EN] Note that FITS does not allow UTF-8, so if we do it might not be possible to translate an arbitrary PDS4 character table into FITS. [PPI] PPI sees negative value in converting tables to FITS format.	Planned for Build 10a and beyond	High	Low ***	Low	Low	Low	Low	Low	8
	PDS4 Delimited Table	[EN] See PDS4 Delimited Separated Value (DSV) standard. [PPI] CSV meets the need for table to delimited table conversion.	Planned for Build 10a and beyond	Medium	Low	Medium	Low	High	Low	Low	10
	VOTable	[PPI] Already have tools to perform this transformation.	Planned for Build 10a and beyond	Low	Low	Low	Medium	Low	Low	Low	7
PDS4 DelimitedTable** (Table_Delimited)	CSV	[EN] PDS DSV to selected standard implementation(s) of CSV.	Included in Build 7b								
	FITS	[EN] Note that FITS does not allow UTF-8, so if we do it might not be possible to translate an arbitrary PDS4 character table into FITS. [PPI] PPI sees negative value in converting tables to FITS format.	Planned for Build 10a and beyond	High	Low ***	Low	Low	Low	Low	Low	8
	PDS4 Character Table	[PPI] This conversion makes ASCII data more difficult to work with.	Planned for Build 10a and beyond	High	Low	Medium	Low	Medium	Low	Low	10

Transformation Prioritization cont.

- The prioritization exercise identified the following as higher priority (assigned to builds):
  - Build 8a
    - PDS3 labeled tables to CSV
    - PDS4 label to PDS3 label (ODL)
  - Build 8b
    - PDS3 labeled tables to PDS4 labeled tables
    - PDS4 labeled tables to PDS4 labeled tables (e.g., binary to character)
  - Build 9a
    - Array 2D Map to GeoTIFF
    - Array 3D Spectrum to ENVI Cube
    - Array 2D Image to FITS

Validation Prioritization

- The Validate Tool supports the following as of the Build 7b release:
  - Syntactic and semantic validation via the XML Schema and 350+ Schematron rules pertaining to PDS4 label structure and content.
  - Bundle and Collection referential integrity checking.
- The team was recently tasked with prioritizing the validation requirements.
  - These focus primarily on specific content validation rules, many of which need to be individually written and schedule for release.
- This was facilitated with the following Google doc:
  - https://docs.google.com/spreadsheets/d/18oqtg3DEo2KrgvBOWLSOuqF2uZtq2XmByJwUknYSZUQ/edit#gid=1268101288

PDS Validate Tool Support

Validation Type	Support
Syntactic Validation	Ensures accuracy of the grammar / syntax of the XML label against the appropriate schemas as well as any mission schema or schematron. *Syntactic validation is largely complete.*
Semantic Validation	Ensures accuracy of the semantic "meaning" of the XML label when run against all relevant schema and schematron files. *PDS has about 350 schematron rules today that are in place.*
Content Validation	**Ensure accuracy of data against the description of the data in the label. *Some content validation in place, but will be the focus of expanded support.***
Referential Integrity	Ensures accuracy of relationships described in Bundle and Collection products *in place today*. Future Feature – Ensure that the relationships described, in and between observational and document objects described in the XML label, are consistent and represented.

Req Number	Requirement Text	Status	Comments	Build	Priority (High, Medium, Low)							
					ATMOS	GEO	IMG	NAIF	PPI	RMS	SBN	PSA
General Validation Requirements												
L5.PRP.VA.1	The tool shall accept the following as input for specifying the product(s) to be validated: a) File specification(s) b) Directory specification(s) c) Web URL(s) in form of a WebDAV end point	Complete		Included in Build 7b	High							Low
L5.PRP.VA.2	The tool shall traverse a directory tree and validate products discovered within that tree.	Complete		Included in Build 7b	High							N/A
L5.PRP.VA.3	The tool shall validate aggregate products and all products referenced by such products.	Complete		Included in Build 7b	Medium							High
L5.PRP.VA.4	The tool shall accept the following as input for specifying the associated schema file(s): a)File specification(s) b)Directory specification(s) c)Web URL(s) d)XML Catalog file e)An xsi:schemaLocation attribute within the product label	Partially Complete Item b) not yet supported. Item d) may require further development (requires review).		Planned for Build 8b	Medium	Medium						Low
L5.PRP.VA.5	The tool shall verify that a schema file is valid.	Complete		Included in Build 7b	High							N/A
L5.PRP.VA.6	The tool shall verify that complex elements defined in a schema file conform to the following naming rules: a)Names must be constructed from one or more components, each of which is made from the ASCII alphanumeric character set. b)Components of the name must be separated by underscores. c)Each component must begin with an upper case letter with all other letters in upper or lower case. d)Articles and prepositions must not be used as components. e)The name must not exceed 255 characters.	Not Addressed Yet		Planned for Build 9a	Medium	Medium						Medium
L5.PRP.VA.7	The tool shall verify that simple elements defined in a schema file conform to the following naming rules: a)Names must be constructed from one or more components, each of which is made from lower-case ASCII letters and/or digits. b)Components of the name must be separated by underscores. c)Each component must begin with a lower case letter with all other letters in lower case. d)Articles and prepositions must not be used as components. e)The name must not exceed 255 characters.	Not Addressed Yet		Planned for Build 9a	Medium	Medium						Medium
L5.PRP.VA.8	The tool shall indicate the schema(s) utilized during validation.	Complete		Included in Build 7b	Low							N/A
L5.PRP.VA.9	The tool shall verify the following file naming rules: a)Label files have an extension of “xml”. (3 [7]) b)File names follow file-naming rules. (6C.1.1 [7]) c)File names do not match prohibited file names. (6C.1.2 [7]) d)File names do not contain prohibited base names. (6C. 1.4 [7]) e)File name extensions are appropriate to the type of product. (6C.1.6[7])	Partially Complete Item e) not supported yet.		Planned for Build 9a	Medium	Medium						High

Validation Prioritization cont.

- Upcoming priorities (from TWG discussions):
  - Build 8a
    - Data content validation of tables against the label description.
  - Build 8b
    - Data content validation of arrays against the label description.
    - Additional integrity checks and lower priority requirements.
- Will revisit once all Nodes have provided input.

Next Steps

- Complete validation requirement prioritization.
- Prioritize requirements and releases for Inspect Tool.
- Prioritize node tools for discussion.
- Continue generating tool “wish list” (need heavy input from the nodes)
- Explore software deployment options including container technology
- Define requirements for GUI front-ends (Validate)
- PDS3 maintenance plans

Questions/Comments

Types of PDS4 Validation

Validation Type	Definition
Syntactic Validation	Validate the grammatical structure / syntax of the XML label
Semantic Validation	Validate that the interpretation of the values in the XML labels are consistent with the circumstances and contexts being described
Content Validation	Validate that the syntactic and semantic content of the XML label accurately describes the digital bits in the data objects being described (i.e., the physical structure and order of the bits are accurately described by the data objects in the xml label)
Referential Integrity	Validate that the relationships described, in and between digital objects described in the XML label, are consistent and represented