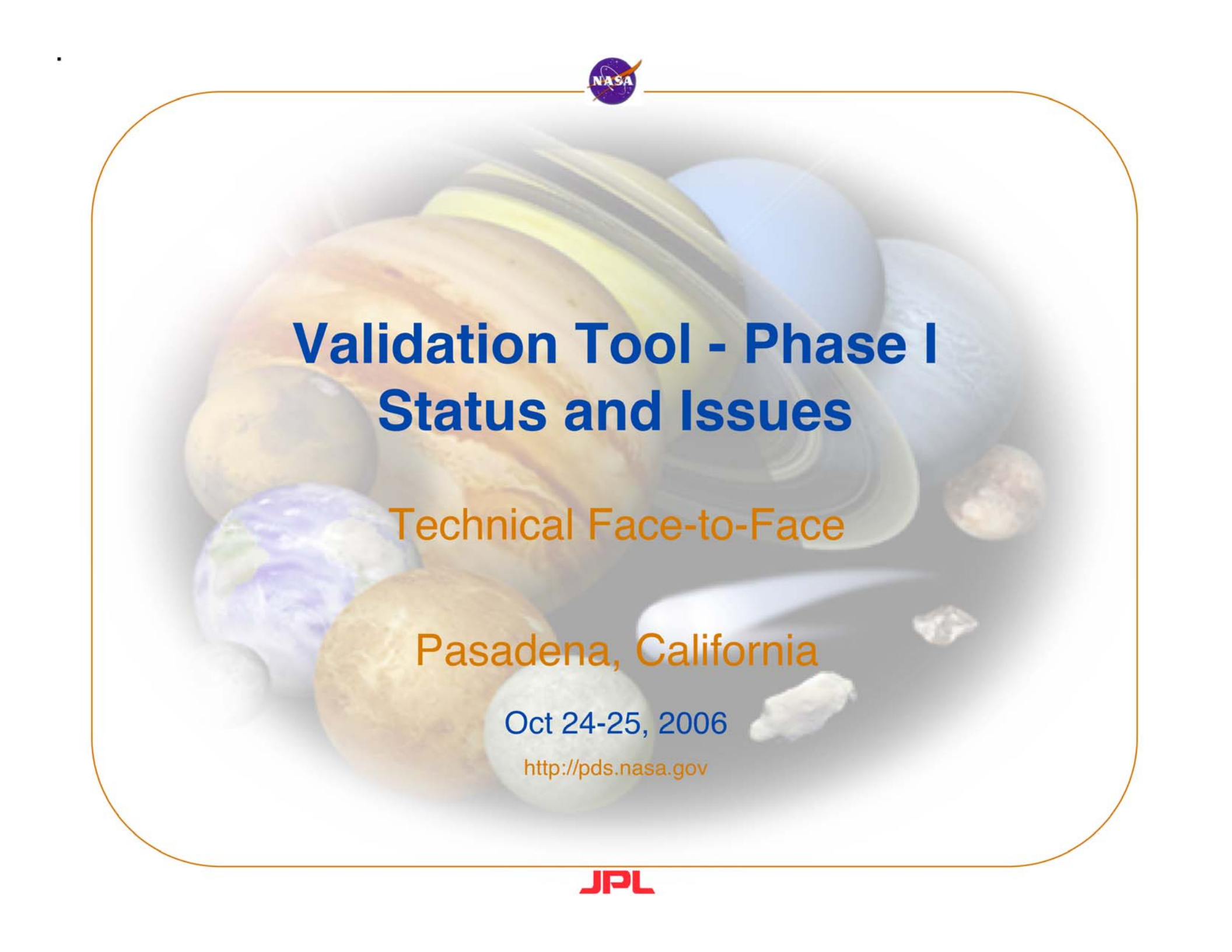

Validation Tool - Phase I Status and Issues

Technical Face-to-Face

Pasadena, California

Oct 24-25, 2006

http://pds.nasa.gov

JPL

Agenda - Tuesday October 24

Check-in Visitor Control	8:30 AM
Introduction, logistics, agenda review	9:00 AM – 9:15 AM
VTool - Phase I status, standards issues, etc for planning Phase II	**9:15 AM – 10:00 AM**
BREAK	10:00 AM – 10:15 AM
VTool - Phase II use case/capabilities discussion	10:15 AM – 11:30 AM
LUNCH	11:30 AM – 1:00 PM
VTool - Phase II requirement identification and discussion	1:00 PM – 2:30 PM
VTool - Phase II summary actions, timeline and release scoping	2:30 PM – 3:15 PM
BREAK	3:15 PM – 3:30 PM
Label design tool - Use case/capabilities discussion	3:30 PM - 5:00 PM
Adjourn	5:00 PM

Topics

- Milestones
- Alpha Test Phase II Results
- Regression Test Suite Status
- Requirements Status
- Validation Tool Functionality (per requirements)
  - Implemented
  - Planned
  - Postponed
- Issues
  - Resolved and Open Tool Issues
  - Resolved and Open Standards Issues
- Schedule

Milestones

- Requirements Review (Feb 2006)
  - Reviewed version 1.1 of the Requirements Document resulting in 66 Requests for Action.
- Released version 1.2 of the Requirements Document (Apr 2006)
  - Resulted in 22 more RFAs.
- Requirements Status and Preliminary Design Telecon (Jun 2006)
  - Issues involving the requirements were presented.
  - The meeting resulted in 11 action items.
- External Code Walk-through (Jun 2006)
  - Several general comments and suggestions were offered by the review board as well as over 50 specific comments regarding the source code.
- Status and Beta-Test Process Telecon (Jul 2006)
  - Presented the test approach and plan for the renamed Alpha Test.
  - The meeting resulted in 8 action items.
- Released version 1.3 of the Requirements Document (Jul 2006)
  - Waiting for comments.
- Alpha Test Phase I with wrap-up Telecon (Aug 2006)
  - Presented the results and resolved several issues.
- Alpha Test Phase II (Sep 2006)
  - Report released last week.

Alpha Test Phase II Results

- Participating Nodes: Atmospheres, Geosciences, Imaging, Rings and Small Bodies.
- Testing of Regression Test Suite
  - Platforms:
    - Solaris 10 (Java 1.5) - Successful execution
    - Windows XP (Java 1.5) - Successful execution at three Nodes
    - Solaris 8 (Java 1.2.2) - Not supported
    - Linux Red Hat AS (GCJ Java 1.4.2) - Unable to complete test
    - Linux Fedora Core 4 (GCJ Java 1.4.2) - Unable to complete test
  - There appear to be some serious issues running the tool under GNU Compiler for Java (GCJ). The EN is currently investigating the problem.
- Testing of Node specific test labels/volumes
  - Some problems reported due to yet to be implemented features.
  - Several requests for documentation clarification and enhancement.
  - Several requests for inclusion of the reporting features.
  - These will be corrected in the next release.

Regression Test Suite Status

- The suite consists of 64 test cases covering the requirements planned for deployment in the initial release.
  - The test cases are comprised of 308 distinct labels.
  - This covers functionality that hasn't been implemented yet, hence all of the errors.
- The suite comes with its own Java application (VTTest) for exercising the test cases with VTool.
  - Functionality planned for this application includes the ability to compare a given test run with the baseline test run provided in the distribution.

Requirements Status

- Version 1.3 of the Validation Tool Requirements document was distributed to the PDS Tech Staff on July 25, 2006.
- RFA Status Summary (88 Total)
  - 16 Open — Related to data object validation.
  - 39 Addressed — Incorporated into version 1.3 of the document.
  - 03 Tabled — Issues to be addressed at a later date.
  - 30 Closed — Finally put to bed.
- Status Legend
  - Open — An action has yet to be defined.
  - Addressed — An action for the RFA has been defined but not yet accepted by the author.
  - Tabled — Waiting on some other milestone before the RFA can be addressed.
  - Closed — The action defined for the RFA was accepted by the author.
- Comments received from Geosciences, Rings and Small Bodies so far.

Implemented Functionality

- General
  - Validation of one or more products, which may reside in a directory, which may reside in a directory tree.
  - Merge the contents of label fragments with the contents of the parent label if the fragments are collocated.
- Syntactic Validation
  - Validation of labels based on the Standards Reference (~95% complete).
  - Verify valid character set and lines terminated with <CR><LF>.
  - Verify date/time values based on the Standards Reference.
- Semantic Validation
  - Validation of labels according to the constructs specified in PDS compliant dictionary (PSDD, local dictionary).
  - Recognize and report the existence of SFDUs in a label.
  - Recognize labels based on existence of PDS_VERSION_ID or SFDU on the first line.
- Initial reporting of validation results.
- Tool is accessible via command-line and API with configurable parameters.
- Initial documentation supporting use of the tool.

Planned Functionality (1 of 3)

Requirement	Aspect	Still To Do
L5.VAL.FR.1	Syntactic Validation	We believe that about 95% of the capabilities are supported. Still need to implement the white space stripping rules for standard value checking.
L5.VAL.FR.2	Semantic Validation	Range checking of values for NON_DECIMAL elements.
L5.VAL.FR.3	Object Existence	Verify the existence of referenced data objects.
L5.VAL.FR.12	Report Content	Need to finalize the content and format of a given message as well as appropriate severity levels.
L5.VAL.FR.13	Report Format	Need to finalize and implement software-readable format.
L5.VAL.FR.14	Report Format	Need to define and implement report views.
L5.VAL.FR.18	Report Format	Need to finalize and implement human-readable format.
L5.VAL.FR.19	File Existence	Verify the existence of referenced files.
L5.VAL.FR.20	Report Content	Need to include tool information, date of execution and configurable parameter settings in the report.

Planned Functionality (2 of 3)

Requirement	Aspect	Still To Do
L5.VAL.FR.23	Label Fragments	Need to implement discovery of label fragments that are not collocated with the parent label.
L5.VAL.FR.27	Label Fragments	Need to implement rules for standalone label fragment validation.
L5.VAL.FR.28	Label Fragments	Need to implement rules for identifying label fragments.
L5.VAL.FR.31	Status	Need to determine what the exit status should represent.
L5.VAL.FR.32	URL Access	Need to fully support URLs for specifying validation targets.
L5.VAL.FR.33	URL Access	Need to fully support URLs for specifying the PSDD and local data dictionaries.

Planned Functionality (3 of 3)

Requirement	Aspect	Still To Do
L5.VAL.FR.36	Semantic Validation	Need to check for required file characteristic elements.
L5.VAL.FR.37	Semantic Validation	Need to check for proper label padding.
L5.VAL.NF.2	Command-Line	Not all capabilities or parameters are supported from the command-line yet.
L5.VAL.NF.6	Parameters	Not all planned parameters are supported yet. See the backup slides for more detail.
L5.VAL.NF.8	Platform	The tool runs on the PDS supported platforms. Still need to determine the problem with GCJ on Linux.
L5.VAL.NF.9	API Docs	Need to enhance and complete the API documentation and provide a Developer Guide.
L5.VAL.NF.10	User Docs	Need to enhance and complete the User Guide.
L5.VAL.NF.11	Report Output	Need to clarify what output goes to standard out versus standard error.

Postponed Functionality

Requirement	Aspect	Still To Do
L5.VAL.FR.9	Data Object Validation	Need to define and implement TABLE object validation.
L5.VAL.FR.10	Data Object Validation	Need to define and implement SPREADSHEET object validation.
L5.VAL.FR.11	Data Object Validation	Need to define and implement IMAGE object validation.
L5.VAL.FR.26	SFDU	Need to validate the content of the SFDU labels.

- Data object validation will be covered in the Phase II discussion.

Resolved Tool Issues

- White Space Stripping
  - Rules for stripping white space from a string-based value, for the purposes of comparing against a standard value list. Necessary for:
    - Values that span multiple lines.
    - Values containing underscores, extraneous spaces or mixed case.
  - The rules are detailed in a backup slide.
- Quoting Values
  - Values containing special characters must be quoted (e.g. “N/A”).
  - Identifier values may be quoted whether it is necessary or not.
  - Values for elements of type integer, time, etc. may not be quoted.
- Alias Support
  - Support for aliases will be disabled by default. Enabled by configurable parameter.
  - When disabled, any occurrence of an element or object named with an alias will result in an element or object not found message, respectively.
  - When enabled, any occurrence of an element or object named with an alias will result in an element or object name has been deprecated message, respectively.

Open Tool Issues

- Report Content and Format
  - Too much detail to cover here.
  - Planning a telecon in early November to discuss reporting.
- GNU Compiler for Java (GCJ)
  - Need to figure out what's up with this.
  - Currently investigating.
- Max Errors Command-Line Option
  - Current default is 300 error messages per run. Carry-over from LVTool.
  - We have a couple of requests from Nodes for this number to represent maximum error messages per label.

Resolved Standards Issues

- PARAMETERS Group redefined from specific object to generic object and replaced the reserved word GROUP with GENERIC_GROUP.
- Unit values must be quoted, single for now, discuss double later.
- Cleaned up several element definitions with regard to the MINIMUM* and MAXIMUM* attribute values:
  - DOWNSAMPLE_METHOD
  - HOUSEKEEPING_CLOCK_COUNT
  - INST_CMPRS_QUANTZ_TYPE
  - MD5_CHECKSUM
  - FIELDS
  - FIELD_NUMBER
- Corrected DATA_SET_TERSE_DESC element definition by removing the bogus standard value.
- Changed X-, Y-, Z-OFFSET keywords to X_, Y_, and Z_OFFSET.
- Standard values for AXIS_NAME and BAND_SEQUENCE must be quoted. Future PSDD to permit STANDARD_VALUE_SETS.
- Decimal points are required in exponential numbers. Corrected the examples in the standards reference.
- Clarification of use of partial seconds and “Z” in time values in Standards Reference.

Open Standards Issues (1 of 5)

- SCR 3-1070: ODL Lines and Records
- SCR 3-1073: Update ARCHIVE_STATUS Keyword
- SCR 3-1079: Update format of pdsdd.full and pdsdd.idx
- SCR 3-1087: PDS Standards on Sequences and Sets
- SCR 3-1089: Add Sub-Objects to FILE Object in PSDD
- SCR 3-1090: StdRef Discrepancy Regarding Signed Non-Decimal Numbers
- SCR 3-1091: Add PPI MISSION_PHASE_NAMEs
- Remove optional PSDD from many OBJECTs
- Bad characters in unit ids
- Handling of aliases
- AMBIENT_TEMPERATURE

Open Standards Issues (2 of 5)

- SCR 3-1070: ODL Lines and Records
  - The rules for handling record boundaries in ODL labels as specified in section 12.4.1 of the Standards Reference are different from PDS common practice. The VTool programmers need to know what standard to program to.
  - Status: Issue discussed at July 26 Standards Telecon. Agreement reached on changed wording for StdRef to reflect current practice; final version of SCR needs to be sent out for formal vote.

- SCR 3-1073: Update ARCHIVE_STATUS Keyword
  - Hyphenated standard values such as “ARCHIVED-ACCUMULATING”, which are not explicitly listed in the PSDD, cause errors in VTool.
  - Status: An SCR has been submitted to fully define and list each standard value. The SCR has not yet been discussed.

Open Standards Issues (3 of 5)

- SCR 3-1079: Update format of pdsdd.full and pdsdd.idx
  - The current format of the pdsdd.full and pdsdd_*.idx files is not consistent with proper ODL syntax, and therefore caused some parsing problems for VTool.
  - Status: The EN Dev team has agreed to postpone resolution of this issue until PDS4. In the meantime, changes have been made to VTool to accommodate the current format of the PSDD files.

- SCR 3-1087: PDS Standards on Sequences and Sets
  - Several ambiguities or inconsistencies regarding the syntax for sequences and sets are present in the Standards Reference. These need to be resolved in order to be properly validated by VTool.
  - Status: This SCR was discussed at the September 20 Standards Telecon. Consensus was reached; a final version of the SCR needs to be written and sent out for a vote.

- SCR 3-1089: Add Sub-Objects to FILE Object in PSDD
  - Neither the PSDD nor the StdRef lists any optional or required sub-objects for the FILE object. Causes errors in VTool.
  - Status: SCR has just been submitted.

Open Standards Issues (4 of 5)

- SCR 3-1090: StdRef Discrepancy Regarding Signed Non-Decimal Numbers
  - Two different places in the StdRef offer differing rules on the inclusion of a sign within non-decimal numbers. This discrepancy needs to be resolved for proper validation by VTool.
  - Status: SCR has just been submitted.

- SCR 3-1091: Add PPI MISSION_PHASE_NAMEs
  - PPI validation tests using VTool produced errors for missing standard values for the MISSION_PHASE_NAME element. These standard values should already have been added to the PSDD.
  - Status: SCR has just been submitted to add these values.

- Remove optional PSDD from many OBJECTs
  - Many PDS objects list “PSDD” as a optional keyword. In some places, this is appropriate, but in others it isn’t. It precludes useful validation of the contents of those objects and its use should be reviewed.
  - Status: SCR has not yet been submitted.

Open Standards Issues (5 of 5)

- Bad characters in unit ids
  - Many unit ids contain characters that are invalid in an ODL context.
  - Status: This issue will be difficult to resolve and the consensus has been to postpone it to PDS4.

- Handling of aliases
  - There are several places within the PSDD where keywords are aliased to each other but have differing standard value lists.
  - Status: The decision has been made to discourage aliases in the current version of VTool. An SCR needs to be submitted to resolve the alias issue.

- AMBIENT_TEMPERATURE
  - MAXIMUM_LENGTH attribute should be set to N/A.
  - Status: Not yet acted upon.

Schedule

- Report Content and Format Telecon (Nov 2006)
  - Present proposed content and format for validation reports.
- Beta Test (Dec 4 - Dec 15, 2006)
  - Open to all Nodes.
- Post-Beta Test Telecon (The week of Dec 18, 2006)
  - Present the results of the last phase of testing.
  - Address and determine the readiness for operational delivery.
- Operational Delivery of version 1.0.0 (Jan 2007)

Backup Slides

White Space Stripping Rules

- Reference PDS Standards (sections 12.5.3.1 and 12.5.4.1).
  - Both sections contain statements and examples relating to this behavior.
- Proposed VTool Behavior
  - Replace <CR><LF> and surrounding spaces with a single space character.
    - Exception: If <CR><LF> is preceded by a hyphen, the hyphen and <CR><LF> are removed without the single space replacement.
  - Compare the value against standard value list. If no match, continue.
  - Remove leading or trailing spaces.
  - Replace underscores with spaces.
  - Replace multiple space characters with a single space.
  - Replace lower-case characters with upper-case characters.
  - Compare the value against the standard value list.
    - Match - Generate warning message regarding manipulation of the value.
    - No Match - Generate normal message regarding value not found.

Command-Line Options (1 of 3)

- Validation
  - **Target (-t, --target)**
    - Specify the file(s) or directories to be validated. May specify as URLs.
    - Required.
  - **Pattern (-p, --pattern)**
    - Specify the file pattern(s) to be matched in the specified directories.
  - **Recursive (-l, --local)**
    - A flag indicating that specified directories should not be traversed recursively.
    - Default is yes.
  - **Ignore Directories (-D, --ignore-dir)**
    - Specify directories and/or directory patterns to ignore.
  - **Ignore Files (-X, --ignore-files)**
    - Specify files and/or file patterns to ignore.
  - **Include Directory (-I, --include)**
    - Specify a path to search for pointer (label fragment) files.
    - Default is the current directory.
  - **Follow Pointers (-F, --no-follow)**
    - A flag indicating that files referenced by ^STRUCTURE pointers will not be verified for existence or included for parent label validation.
    - Default is yes.
  - **Label Fragments (-f, --fragment)**
    - A flag indicating that standalone label fragments will be validated.
    - Default is no.

Command-Line Options (2 of 3)

- Validation (cont)
  - **Dictionary (-d, --dict)**
    - Specify the PDS compliant data dictionary(s) to be referenced for validation. Assumes the full version and not the index.
    - If not provided, dictionary (semantic) validation is not performed.
  - **Aliases (-a, --alias)**
    - A flag indicating that aliases from the specified data dictionaries are enabled.
    - Default is no.
  - **Data Object Validation (-OBJ, --no-obj)**
    - A flag indicating that data objects will not be validated.
    - Default is yes.
- Miscellaneous
  - **Configuration (-c, --config)**
    - Specify a file containing configuration parameters.
  - **Help (-h, --help)**
    - A flag indicating that the application usage (command-line options) should be displayed to the terminal.
  - **Version (-V, --version)**
    - A flag indicating that the application version should be displayed to the terminal.

Command-Line Options (3 of 3)

- Reporting
  - **Report (-o, --output)**
    - Specify the file for the report.
    - Default is to send the report to the terminal.
  - **Detail (-od, --output-detail <full|sum|min>)**
    - Specify the report detail (Full, Summary or Minimal).
    - Default is full.
  - **Severity (-v, --verbose <0|1|2|3>)**
    - Specify the message severity level and above to include (0=Debug, 1=Info, 2=Warn, 3=Error and Fatal).
    - Default is Warn (2).
  - **XML Format (-x, --xml-output)**
    - A flag indicating that the report is output in machine-readable (XML) format.
    - Default is human-readable.
  - **Max Errors (-m, --max-errors)**
    - Specify the maximum number of errors to report for a validation run.
    - Default is 300.
  - **Progress Report (-pr, --progress)**
    - A flag indicating that validation progress will be reported to the terminal.
    - Default is no.