



Validation In-Depth

PDS Technical Session
Michael Cayan

February 13, 2018



Topics



- Features
- Validation Scenarios
- Demonstration
- Wrap Up



Features

- Validation against a Schema and Schematron through the following means:
 - Core
 - User-specified
 - The label itself
 - An XML Catalog file
- Validation against File References in the Product Label
 - Check that file references do exist.
 - Casing of file references match their physical files on the file system.
 - Checksum validation against the actual checksum and a Manifest file (if supplied).
- Validation at the Bundle level, including, but not limited to:
 - Referential integrity.
 - Names of files and directories adhere to the PDS4 Standards.
- Validation of Tables



Table Content Validation Features

- For all tables, field values are checked for the following:
 - Values match their defined data type.
 - Values match their defined field format, if specified.
 - Values are within the defined min and max values of the field, if specified.
- Specifically for binary tables, fields are checked for the following:
 - Fields that are packed match the defined number of bit-fields.
 - Bit-field values that match their defined data type.
- Specifically for character and delimited tables, records are checked for the following:
 - Records in the data file end in CRLF.
 - The length of a record does not exceed the defined length (or defined maximum length in the case of delimited tables).
 - Number of records match the defined number in the label.



Validation Scenarios

- Single product
 - Spot checking or in a pipeline
- Directory of products
 - No specific structure
- Bundle/Collection of products
 - Incorporates referential integrity checks
- Schema referencing options
 - Supplied with the tool
 - As referenced in the label
 - As referenced in an XML Catalog file



Demonstration

Live demonstration will be of the following scenarios:

- Validating a directory of products
- Validating with an XML Catalog
- Validating a bundle
- Validating products in a pipeline



Validating a Directory of Products

- Validate against the latest schema and schematron packaged with the tool:

```
%> validate /home/dph_example_archive
```

- `-f` flag option validates against the schemas and schematrons defined in the product label:

```
%> validate /home/dph_example_archive -f
```



Validation with an XML Catalog



Pass in the Catalog file with the -C flag:

```
%> validate /home/dph_example_archive -C catalog.xml
```




Validating a Bundle

- Specify the *pds4.bundle* rule as well as a Checksum Manifest file to do validation of a bundle:

```
%> validate /home/dph_example_archive -R pds4.bundle -M manifest.md5
```

- If the *-R pds4.bundle* rules flag is not specified, the tool will perform a file by file validation run.
- If the Checksum Manifest file contains relative file references, then the tool will assume that the base path to these references is the target directory.



Validating Products in a Pipeline

In a pipeline scenario, the tool is usually run against a single product, where the output report is in either XML or JSON:

- For an XML-style report:

```
%> validate product.xml -s xml
```

or

- For a JSON-style report:

```
%> validate product.xml -s json
```



Wrap Up

- Current plans have development of prioritized requirements through Builds 10a and 10b
- Feature enhancements may also be requested and will be prioritized by the TWG as well

Questions/Comments