

PDS4 Label Design Tutorial—Validation

VIPER DAWG, JUNE 23, 2022

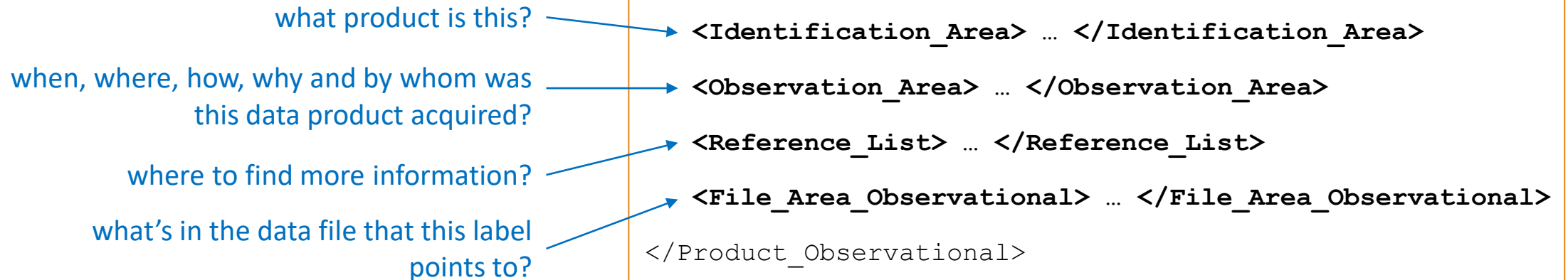
SUSIE SLAVNEY, PDS GEOSCIENCES NODE

A solid orange horizontal bar at the bottom of the slide.

Outline – where did we leave off?

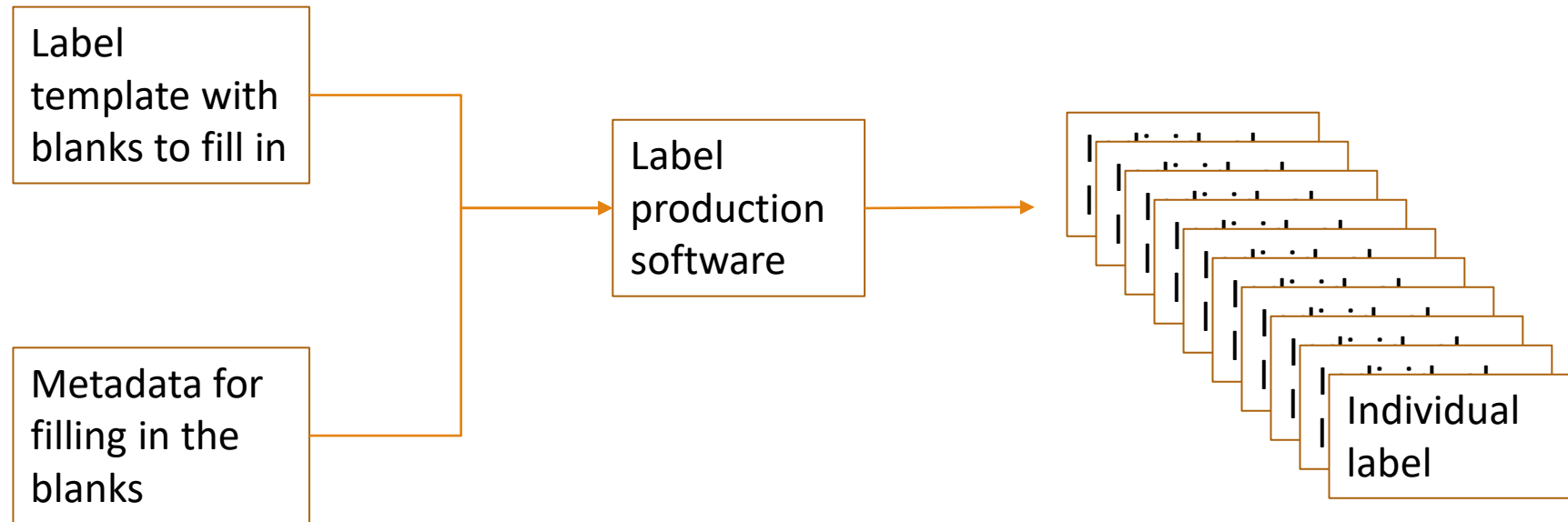
1. What a PDS Label Is For
2. Anatomy of a PDS4 Label
3. Examples Used In This Tutorial
4. How To Design a Label
 - a. Identification_Area
 - b. File_Area_Observational
 - c. Reference_List
 - d. Observation_Area
5. What To Do With the Label You Have Designed
 - a. You have a label design, so now what?
 - b. Turning a draft label into a label template
 - c. Turning a label template into a batch of labels
 - d. **Validating labels**
6. References

How To Design A Label—Recap



You have a label design, so now what?

How to go from label design to production of real labels?



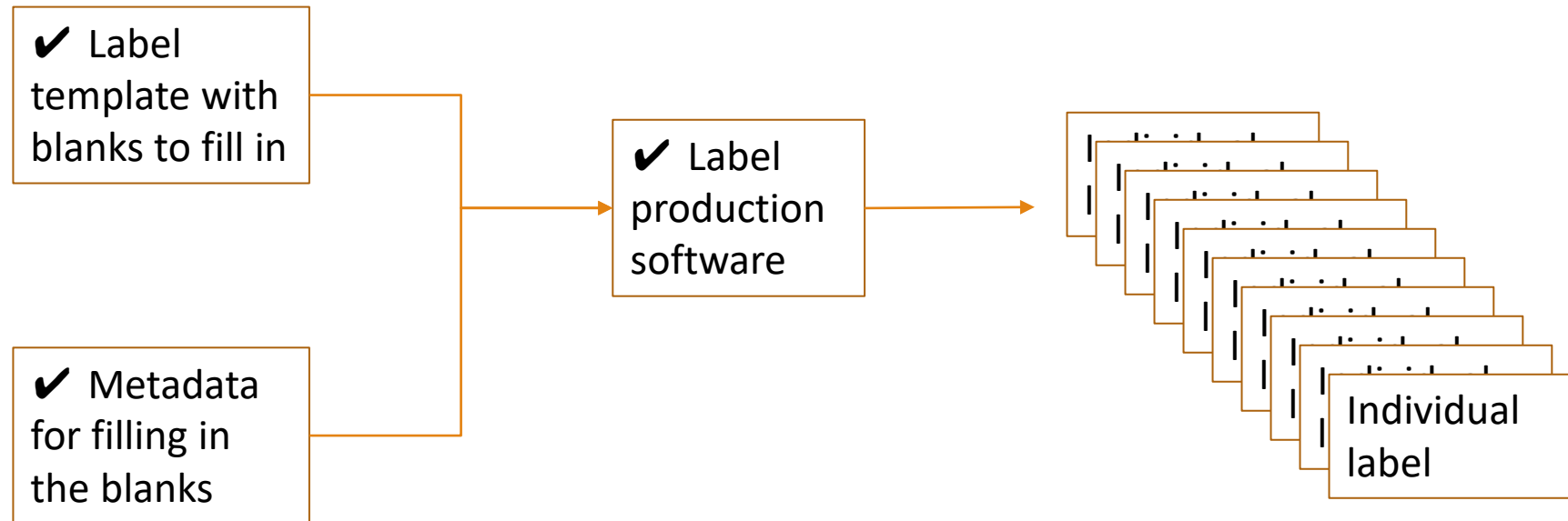
Turning a Draft Label into a Label Template

Replace the values that vary with placeholders that the label processing software understands.

```
<Product_Observational ... >
  <Identification_Area>
    <logical_identifier>urn:nasa:pds:viper_nirvss_spc:data_raw:${product-name-lowercase}</logical_identifier>
    <version_id>1.0</version_id>
    <title>VIPER NIRVSS Data Product ${product-name}</title>
    ...
  </Identification_Area>
  <Observation_Area>
    <Time_Coordinates>
      <start_date_time>${start-time-utc}</start_date_time>
      <stop_date_time>${stop-time-utc}</stop_date_time>
    </Time_Coordinates>
    ...
  </Observation_Area>
  ...
  <File_Area_Observational>
    <File>
      <file_name>${data-file-name}</file_name>
      <local_identifier>${data-file-name-minus-extension}</local_identifier>
      <creation_date_time>${data-file-creation-time}</creation_date_time>
    </File>
    <Table_Delimited> ... </Table_Delimited>
  </File_Area_Observational>
</Product_Observational>
```

PDS MI-Label and Docgen use the Velocity template language with placeholders that look like `${this}`, but they can be much more complex

Turning a Label Template into a Batch of Labels



Outline – one more step

1. What a PDS Label Is For
2. Anatomy of a PDS4 Label
3. Examples Used In This Tutorial
4. How To Design a Label
 - a. Identification_Area
 - b. File_Area_Observational
 - c. Reference_List
 - d. Observation_Area
5. What To Do With the Label You Have Designed
 - a. You have a label design, so now what?
 - b. Turning a draft label into a label template
 - c. Turning a label template into a batch of labels
 - d. **Validating labels**
6. References

Validation Steps

Why validate labels?

Downloading and installing Validate

Downloading and installing dictionaries, if you need local copies

How to run the PDS Validate Tool

Reading the validation report

How to run the Validate Tool using local copies of dictionaries

Validate command-line switches

What to do if you think you have found a bug

Why Validate Labels?

PDS requires that you run the PDS tool **Validate** on your data products before delivery to PDS.

- This helps you catch and correct errors that would prevent the delivery from being accepted.
- The validation report should be included with the delivery.
- PDS will also run Validate upon receipt of the delivery.

Validate checks many things, including:

- Do the labels correctly describe the data product format and contents?
- Do the metadata conform to the specifications in the PDS main dictionary and any other dictionaries?
- Do the collection and bundle labels correctly list all the products in the delivery?
- Do the labels correctly refer to registered PDS context products?

PDS will help data providers install and use Validate.

Downloading and Installing Validate

The PDS Validate tool is here: <https://nasa-pds.github.io/validate/>.

It requires Java version 9.0 or higher.



Validate Tool

Last Published: 2022-06-09 | Version: 2.2.3

[Validate Github Repo](#) 

click here

DOCUMENTATION

About
Validate Tool

Installation

About Validate Tool

The Validate Tool project contains software for validating PDS4 product labels and product data. The associated specific schema for the product label specifies syntactic and semantic constraints. The product label itself specifies the constraints for the data.

Downloading and Installing Validate

In the GitHub repository

The screenshot shows the GitHub repository page for NASA-PDS/validate. The repository is public and has 3 forks and 9 stars. The file list on the left includes folders like .github, build, docs/requirements, and src, and files like .gitignore, .zenodo.json, CHANGELOG.md, LICENSE.md, NOTICE.txt, README.md, and SECURITY.md. The 'About' section on the right provides information about the repository, including a link to nasa-pds.github.io/validate/ and a list of tags (validation, nasa, pds, pds4, nasa-pds). The 'Releases' section is highlighted with a red circle, showing the latest release 'validate v2.2.3' with a 'Latest' badge and a date of 13 days ago. A red arrow points to this release with the text 'click on Latest'.

Downloading and Installing Validate

The screenshot shows the GitHub repository page for NASA-PDS/validate. The release v2.2.3 is highlighted as the latest version. The assets section lists several files, with 'Documentation (zip)' and 'validate-2.2.3-bin.tar.gz' circled in red. A red text box provides instructions on how to choose between these two assets.

Asset Name	Size	Released
Documentation (zip)	9.77 MB	13 days ago
validate-2.2.3-bin.tar.gz	48 MB	13 days ago
validate-2.2.3-bin.zip	49 MB	13 days ago
validate-2.2.3-javadoc.jar	1.36 MB	13 days ago

Choose either `validate-x.x.x-bin.tar.gz` or `validate-x.x.x-bin.zip`

Downloading and Installing Validate

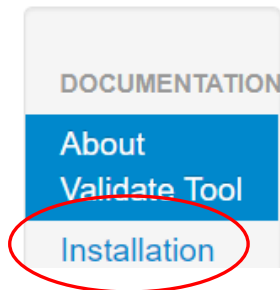
Back on <https://nasa-pds.github.io/validate/>, choose Installation and follow the instructions.



Validate Tool

Last Published: 2022-06-09 | Version: 2.2.3

[Validate Github Repo](#) 



About Validate Tool

The Validate Tool project contains software for validating PDS4 product labels and product data. The associated specific schema for the product label specifies syntactic and semantic constraints. The product label itself specifies the constraints for the data

[click here](#)

Downloading and Installing Dictionaries, If You Need Local Copies

Validate needs access to the PDScore dictionary and any other dictionaries used in your labels.

Published dictionaries are online at <https://pds.nas.gov/datastandards/dictionaries/>. If you have Internet access, Validate will find them online. You don't need to do anything to make this happen.

If your labels use a dictionary that is not yet published, you will need local copies of all the dictionaries mentioned in your labels, not just the unpublished ones.

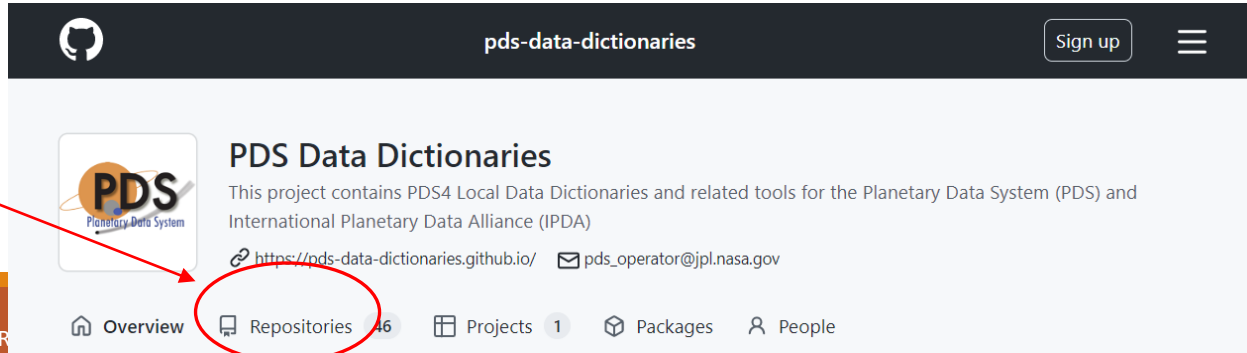
- *Q: Why would I want to use an unpublished dictionary?*

A: Maybe because it's a new mission dictionary that is still a working draft, or maybe it's a revised dictionary that you are testing before it becomes public.

- *Q: So where do I get these unpublished dictionaries?*

A: On <https://github.com/pds-data-dictionaries/>.

click here



The screenshot shows the GitHub repository page for 'pds-data-dictionaries'. The repository name is at the top right. Below it, the repository description reads: 'This project contains PDS4 Local Data Dictionaries and related tools for the Planetary Data System (PDS) and International Planetary Data Alliance (IPDA)'. The repository URL is <https://pds-data-dictionaries.github.io/> and the contact email is pds_operator@jpl.nasa.gov. At the bottom, the navigation bar includes 'Overview', 'Repositories' (which is circled in red and has a red arrow pointing to it from the text 'click here'), 'Projects' (1), 'Packages', and 'People'.

Downloading and Installing Dictionaries, If You Need Local Copies

- *Q: Do you really need local copies of all the dictionaries mentioned in the label? There should be a way to refer to the published dictionaries online, and the unpublished ones locally.*

A: Yes, there should be a way, and it might be possible using a method called **catalog files**, which is a way of tricking the software into looking somewhere else for dictionaries. It's an advanced topic beyond this tutorial. I have not figured out how to use a catalog file to get Validate to look both online and locally. Validate aficionados, here is your chance to improve this tutorial.

Downloading and Installing Dictionaries, If You Need Local Copies

Put local copies of dictionaries into a directory structure like this, to mimic the PDS structure:

```
local directory
|-- pds4
|   |-- pds
|       |-- v1
|           |-- PDS4_PDS_xxxx.*
|
|   |-- geom (or proc or cart or other discipline dictionary)
|       |-- v1
|           |-- PDS4_GEOM_xxxx_yyyy.*
|
|   |-- mission
|       |-- viper
|           |--v1
|               |-- PDS4_VIPER_xxxx_yyyy.*
```


How To Run Validate

Back on <https://nasa-pds.github.io/validate/>, choose Operation for instructions.

PDS Planetary Data System

Last Published: 2022-06-09 | Version: 2.2.3

Validate Github Repo

Validate Tool

DOCUMENTATION

- About Validate Tool
- Installation
- Operation**
- Common Errors
- Support

About Validate Tool

The Validate Tool project contains software for validating PDS4 product labels and product data. The associated specific schema for the product label specifies syntactic and semantic constraints. The product label itself specifies the constraints for the data.

- [Tool Overview](#)
- [Support](#)

How To Run Validate – simple examples

> validate mylabel.xml -r report.txt

- *Validate the label **mylabel.xml** and write the report to the file **report.txt**. Content validation is on by default.*

> validate \path\to\mydirectory -D -r report.txt

- *Validate all labels in and below **mydirectory**, disabling content validation.*

> validate \path\to\collectiondirectory -R pds4.collection -r report.txt

- *Validate all labels in and below **collectiondirectory**, and do additional collection checks.*

> validate \path\to\bundledirectory -R pds4.bundle -r report.txt

- *Validate all labels in and below **bundledirectory**, and do additional bundle and collection checks.*

Reading the Validation Report

PDS Validate Tool Report

Configuration:

Version 2.2.3
Date 2022-06-22T20:32:58Z

Parameters:

Targets [file:/C:/temp/test-label.xml]
Severity Level WARNING
Recurse Directories true
File Filters Used [*.xml, *.XML]
Data Content Validation on
Product Level Validation on
Max Errors 100000
Registered Contexts File C:\Users\s1avney\MY-PRO~1\validate\VALIDA~1.3\bin\..\resources\registered_context_products.json

Product Level Validation Results

FAIL: file:/C:/temp/test-label.xml

ERROR [error.label.context_ref_not_found] line 32: 'Context product not found: urn:nasa:pds:context:investigation:mission.viper
ERROR [error.label.context_ref_not_found] line 41: 'Context product not found: urn:nasa:pds:context:instrument_host:spacecraft.viper
ERROR [error.label.context_ref_not_found] line 49: 'Context product not found: urn:nasa:pds:context:instrument:mars2020.nirrvss

Begin Content Validation: file:/C:/temp/test-label.csv

ERROR [error.table.fields_mismatch] table 1, record 1: Record 1 has wrong number of fields (expected 1216, got 1217)
ERROR [error.table.field_value_data_type_mismatch] table 2, record 1, field 17: Value does not match its data type 'ASCII_Integer': '-35.74632263183594' do
ERROR [error.table.field_value_data_type_mismatch] table 2, record 2, field 17: Value does not match its data type 'ASCII_Integer': '-38.24736022949219' do

First, Skip To The End...

```
ERROR [error.table.field_value_data_type_mismatch] table 2, record 20, field 17: Value does not match its data type 'ASCII_Integer': '-40.21790313720703'  
does not match the pattern '[+-]?\d+'  
End Content Validation: file:/C:/temp/test-label.csv  
1 product validation(s) completed
```

Summary:

```
24 error(s)  
0 warning(s)
```

Product Validation Summary:

```
0 product(s) passed  
1 product(s) failed  
0 product(s) skipped
```

Referential Integrity Check Summary:

```
0 check(s) passed  
0 check(s) failed  
0 check(s) skipped
```

Message Types:

```
20 error.table.field_value_data_type_mismatch  
3 error.label.context_ref_not_found  
1 error.table.fields_mismatch
```

End of Report

Search For The Error Codes

```
Product Level Validation Results
error.table.fields_mismatch
FAIL: file:/C:/temp/test-label.xml
ERROR [error.label.context_ref_not_found] line 32: 'Context product not found: urn:nasa:pds:context:investigation:mission.viper
ERROR [error.label.context_ref_not_found] line 41: 'Context product not found: urn:nasa:pds:context:instrument_host:spacecraft.viper
ERROR [error.label.context_ref_not_found] line 49: 'Context product not found: urn:nasa:pds:context:instrument:mars2020.nirrvss
Begin Content Validation: file:/C:/temp/test-label.csv
ERROR [error.table.fields_mismatch] table 1, record 1: Record 1 has wrong number of fields (expected 1216, got 1217)
ERROR [error.table.field_value_data_type_mismatch] table 2, record 1, field 17: Value does not match its data type 'ASCII_Integer': '-35.74632263183594' do
ERROR [error.table.field_value_data_type_mismatch] table 2, record 2, field 17: Value does not match its data type 'ASCII_Integer': '-38.24736022949219' do
ERROR [error.table.field_value_data_type_mismatch] table 2, record 3, field 17: Value does not match its data type 'ASCII_Integer': '-39.53461456298828' do
ERROR [error.table.field_value_data_type_mismatch] table 2, record 4, field 17: Value does not match its data type 'ASCII_Integer': '-40.1835823059082' doe
ERROR [error.table.field_value_data_type_mismatch] table 2, record 5, field 17: Value does not match its data type 'ASCII_Integer': '-40.56357192993164' do
ERROR [error.table.field_value_data_type_mismatch] table 2, record 6, field 17: Value does not match its data type 'ASCII_Integer': '-40.46796417236328' do
ERROR [error.table.field_value_data_type_mismatch] table 2, record 7, field 17: Value does not match its data type 'ASCII_Integer': '-40.39441680908203' do
ERROR [error.table.field_value_data_type_mismatch] table 2, record 8, field 17: Value does not match its data type 'ASCII_Integer': '-40.377254486083984' do
ERROR [error.table.field_value_data_type_mismatch] table 2, record 9, field 17: Value does not match its data type 'ASCII_Integer': '-40.35519027709961' do
ERROR [error.table.field_value_data_type_mismatch] table 2, record 10, field 17: Value does not match its data type 'ASCII_Integer': '-40.357643127441406' c
ERROR [error.table.field_value_data_type_mismatch] table 2, record 11, field 17: Value does not match its data type 'ASCII_Integer': '-40.357643127441406' c
ERROR [error.table.field_value_data_type_mismatch] table 2, record 12, field 17: Value does not match its data type 'ASCII_Integer': '-40.35028839111328' do
ERROR [error.table.field_value_data_type_mismatch] table 2, record 13, field 17: Value does not match its data type 'ASCII_Integer': '-40.32332229614258' do
ERROR [error.table.field_value_data_type_mismatch] table 2, record 14, field 17: Value does not match its data type 'ASCII_Integer': '-40.29635238647461' do
ERROR [error.table.field_value_data_type_mismatch] table 2, record 15, field 17: Value does not match its data type 'ASCII_Integer': '-40.29390335083008' do
ERROR [error.table.field_value_data_type_mismatch] table 2, record 16, field 17: Value does not match its data type 'ASCII_Integer': '-40.26693344116211' do
ERROR [error.table.field_value_data_type_mismatch] table 2, record 17, field 17: Value does not match its data type 'ASCII_Integer': '-40.26203155517578' do
ERROR [error.table.field_value_data_type_mismatch] table 2, record 18, field 17: Value does not match its data type 'ASCII_Integer': '-40.244869232177734' c
ERROR [error.table.field_value_data_type_mismatch] table 2, record 19, field 17: Value does not match its data type 'ASCII_Integer': '-40.22280502319336' do
ERROR [error.table.field_value_data_type_mismatch] table 2, record 20, field 17: Value does not match its data type 'ASCII_Integer': '-40.21790313720703' do
```

How To Run Validate Using Local Copies of Dictionaries

How do you tell Validate where to find local copies of dictionaries?

```
> validate \path\to\mydirectory -r report.txt
-x \PDS-Stuff\pds4\pds\v1\PDS4_PDS_1G00.xsd,
  \PDS-Stuff\pds4\mission\viper\v1\PDS4_VIPER_1G00_1100.xsd
-s \PDS-Stuff\pds4\pds\v1\PDS4_PDS_1G00.sch,
  \PDS-Stuff\pds4\mission\viper\v1\PDS4_VIPER_1G00_1100.sch
```

- *Validate all labels in and below **mydirectory**, using these schema and Schematron files*

Well, that's messy. You don't want to be typing all that stuff every time. There's a better way.

How To Run Validate Using Local Copies of Dictionaries

Using a configuration file:

```
> validate \path\to\mydirectory -c config.txt
```

*Validate all labels in and below **mydirectory**, using settings in **config.txt***

Contents of config.txt:

starts a comment

Unix-style path delimiters

```
# This is a Validate Tool configuration file

validate.schema = /PDS-Stuff/pds4/pds/v1/PDS4_PDS_1G00.xsd, \
                  /PDS-Stuff/pds4/mission/viper/v1/PDS4_VIPER_1G00_1100.xsd
validate.schematron = /PDS-Stuff/pds4/pds/v1/PDS4_PDS_1G00.sch, \
                      /PDS-Stuff/pds4/mission/viper/v1/PDS4_VIPER_1G00_1100.sch

validate.verbose = 1
validate.skipContentValidation = true
validate.report = report.txt
```

\ continues on the next line

How To Run Validate Using Local Copies of Dictionaries

If Validate gives a warning message like this:

WARNING [warning.label.schema] line 8, 109: schema_reference.4: Failed to read schema document 'https://pds.nasa.gov/pds4/pds/v1/PDS4_PDS_1G00.xsd', because 1) could not find the document; 2) the document could not be read; 3) the root element of the document is not <xsd:schema>.

... and/or an error message like this:

ERROR [error.label.schema] line 8, 109: cvc-elt.1.a: Cannot find the declaration of element 'Product_Observational'.

... it is most likely because Validate is not looking in the right place for the dictionary files.

That means you should check the dictionary path and file names given in your configuration file.

Validate Command Line Switches

Validate has many command-line switches. To see the list, type

```
> validate -h
```

They are explained in the online help at <https://nasa-pds.github.io/validate/operate/>.

What To Do If You Think You Have Found A Bug in Validate

- Check with your PDS rep to see if you have overlooked something obvious.
- Update Validate.
- Update local copies of dictionaries and their references in the labels.
- See if the problem has already been reported on GitHub at <https://github.com/NASA-PDS/validate/issues>.
- Report it yourself using the New Issue button on that page.

Outline—almost finished

1. What a PDS Label Is For
2. Anatomy of a PDS4 Label
3. Examples Used In This Tutorial
4. How To Design a Label
 - a. Identification_Area
 - b. File_Area_Observational
 - c. Reference_List
 - d. Observation_Area
5. What To Do With the Label You Have Designed
 - a. You have a label design, so now what?
 - b. Turning a draft label into a label template
 - c. Turning a label template into a batch of labels
 - d. Validating labels
6. **References**

References

PDS Data Provider's Handbook —big-picture guide to preparing an archive	https://pds.nasa.gov/datastandards/documents/dph/current/
Small Bodies Node PDS4 Wiki —very detailed label-making instructions, XML background info	https://sbnwiki.astro.umd.edu/wiki/SBN_PDS4_Wiki
The Mars 2020 PIXL archive , where the example label came from	https://pds-geosciences.wustl.edu/missions/mars2020/pixl.htm
PDS Context Products	https://pds.nasa.gov/data/pds4/context-pds4/
PDS Main, Discipline, and Mission Data Dictionaries on the PDS web site	https://pds.nasa.gov/datastandards/dictionaries/
PDS Dictionaries on GitHub	https://github.com/pds-data-dictionaries/
PDS Validate Tool	https://nasa-pds.github.io/validate/