

PDS4 Label Design Tutorial

CLPS DASS, JUNE 16, 2022

SUSIE SLAVNEY, PDS GEOSCIENCES NODE

Outline

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

What a PDS Label Is For

Unique identifier for a data product

Context of the data product: what instrument? what spacecraft? what mission? or what telescope, observatory, or laboratory?

Metadata, such as: what, when, where, how, why, and by whom were the data collected

Data file contents and format: size, data type, rows and columns or lines and samples

Labels are software-readable and human-readable

Similar products have similar labels

Anatomy of a PDS4 Label

First things first: XML (eXtensible Markup Language)

Like HTML, but generalized to describe other kinds of information

Software-readable and human-readable (really)

Controlled by XML schema; i.e., dictionaries

Classes, attributes, and tags

- A tag is a thing inside angle brackets. Every tag comes with a closing tag.
- An attribute is a simple (lowest-level) tag.
- A class is a group of attributes. Classes may be nested.

Made-up example of XML

```
<Movie>
  <title>Bedtime for Bonzo</title>
  <firstRelease>1951</firstRelease>
  <director>Frederick de Cordova</director>
  <screenplayBy>Lou Breslow</screenplayBy>
  <screenplayBy>Val Burton</screenplayBy>
  <storyBy>Ted Berkman</storyBy>
  <storyBy>Raphael Blau</storyBy>
  <starring>Ronald Reagan</starring>
  <starring>Diana Lynn</starring>
</Movie>
```

Anatomy of a PDS4 Label

XML header (always the same)

XML header (Schematron links;
more about these later)

This is a science product

Product identification

Metadata

References

Data file contents and format

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<?xml-model ...>
<?xml-model ...>
<Product_Observational ... >
  <Identification_Area> ... </Identification_Area>
  <Observation_Area> ... </Observation_Area>
  <Reference_List> ... </Reference_List>
  <File_Area_Observational> ... </File_Area_Observational>
</Product_Observational>
```

Schema links go here;
more about these later

Examples Used In This Tutorial

For this tutorial we will use the label for an actual Mars 2020 PIXL raw data product to illustrate some concepts.

- https://pds-geosciences.wustl.edu/m2020/urn-nasa-pds-mars2020_pixl/data_raw_spectroscopy/sol_00167/PS_0167_0681733043_000ENB_0060000631116812139_J02.xml

We will also use the VIPER mission and instruments as examples to show how to design a new label. The actual VIPER labels are still in development and may not end up looking exactly like these examples.

How To Design A Label, Parts 1 to 4

Part 1: Identification_Area

Part 2: File_Area_Observational

Part 3: Reference_List

Part 4: Observation_Area
data dictionaries
context products

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<?xml-model ...>
<?xml-model ...>

<Product_Observational ... >

    <Identification_Area> ... </Identification_Area>

    <Observation_Area> ... </Observation_Area>

    <Reference_List> ... </Reference_List>

    <File_Area_Observational> ... </File_Area_Observational>

</Product_Observational>
```

The four label areas are required to appear in the order shown in the example label above. We will discuss them in a slightly different order that may make them easier to understand.

How To Design A Label—Part 1, Identification_Area

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<?xml-model ...>
<?xml-model ...>
<Product_Observational ... >
  <Identification_Area> ... </Identification_Area>
  <Observation_Area> ... </Observation_Area>
  <Reference_List> ... </Reference_List>
  <File_Area_Observational> ... </File_Area_Observational>
</Product_Observational>
```

what product is this? →

when, where, how, why and by whom was this data product acquired? →

where to find more information? →

what's in the data file that this label points to? →

Before we go into the Identification_Area there will be a slight digression about LIDs (Logical Identifiers) ...

Digression: LIDs, VIDs, and LIDVIDs

Each PDS4 product has a Logical Identifier (LID) and Version Identifier (VID)

When combined, they make a LIDVID, a unique product identifier:

`urn:nasa:pds:mars2020_pixl:data_raw_spectroscopy:ps__0167_0681733043_000enb__00600000631116812139__j.csv::2.0`

bundle

collection

product

version

```
<Identification_Area>
  <logical_identifier>urn:nasa:pds:mars2020_pixl:data_raw_spectroscopy:ps__0167_0681733043_000enb__00600000631
116812139__j.csv</logical_identifier>
  <version_id>2.0</version_id>
  <title>
    Mars2020 PIXL_SPECT ENB Observational Product - ps__0167_0681733043_000enb__00600000631116812139__j.CSV
  </title>
  ...
</Identification_Area>
```

How To Design A Label—Part 1, Identification_Area

Using an example raw spectrum data file from the VIPER NIRVSS instrument (*switch to editor display of NIRVSS data file and draft label*)

what product is this?

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<?xml-model ...>
<?xml-model ...>
<Product_Observational ... >
  <Identification_Area> ... </Identification_Area>
  <Observation_Area> ... </Observation_Area>
  <Reference_List> ... </Reference_List>
  <File_Area_Observational> ... </File_Area_Observational>
</Product_Observational>
```

How To Design A Label—Part 2, File_Area_Observational

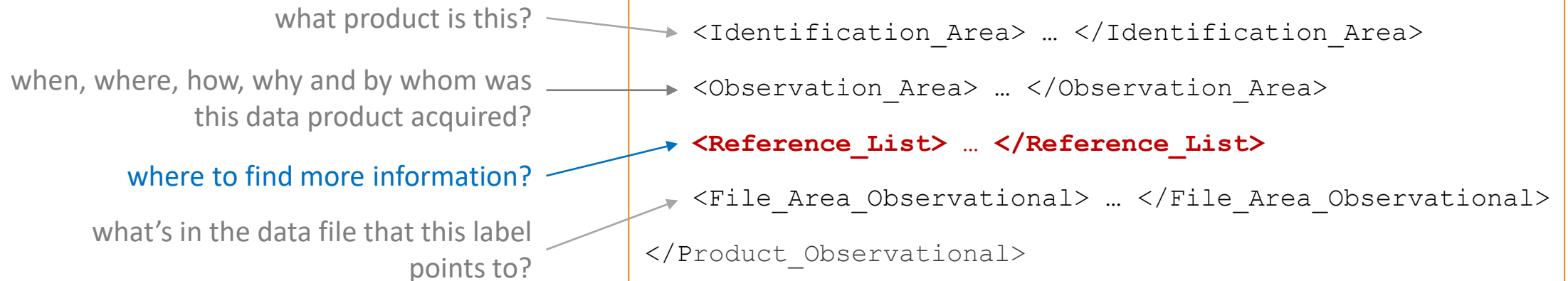
Using an example raw spectrum data file from the VIPER NIRVSS instrument (*switch to editor display of NIRVSS data file and draft label*)

Notice how the spectrum is labeled as a repeating group, not 1200 individual columns (“fields”)

what’s in the data file that this label points to?

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<?xml-model ...>
<?xml-model ...>
<Product_Observational ... >
  <Identification_Area> ... </Identification_Area>
  <Observation_Area> ... </Observation_Area>
  <Reference_List> ... </Reference_List>
  <File_Area_Observational> ... </File_Area_Observational>
</Product_Observational>
```

How to Design a Label—Part 3, Reference_List



How To Design A Label—Part 3, Reference_List

Internal_Reference

- documentation
- source product
- browse product
- anything with a LID

External_Reference

- anything with a DOI

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<?xml-model ...>
<?xml-model ...>
<Product_Observational ... >
  <Identification_Area> ... </Identification_Area>
  <Observation_Area> ... </Observation_Area>
  <Reference_List> ... </Reference_List>
  <File_Area_Observational> ... </File_Area_Observational>
</Product_Observational>
```

Example of Reference_List from the Mars 2020 PIXL bundle label

```
<Reference_List>
...
<Internal_Reference>
  <lid_reference>urn:nasa:pds:mars2020_pixl:document:pixl_rdr_sis</lid_reference>
  <reference_type>bundle_to_document</reference_type>
</Internal_Reference>

<Internal_Reference>
  <lid_reference>urn:nasa:pds:mars2020_pixl:document:pixl_user_guide</lid_reference>
  <reference_type>bundle_to_document</reference_type>
</Internal_Reference>

<External_Reference>
  <doi>10.1007/s11214-020-00767-7</doi>
  <reference_text>
    Allwood, A.C., Wade, L.A., Foote, M.C. et al. PIXL: Planetary Instrument for X-Ray Lithochemistry. Space Sci Rev 216, 134 (2020). https://doi.org/10.1007/s11214-020-00767-7
  </reference_text>
</External_Reference>
...
</Reference_List>
```

How To Design A Label—Part 4, Observation Area

*Return to the editor display of
NIRVSS draft label to view the
Observation_Area*

what product is this?

when, where, how, why and by whom was
this data product acquired?

where to find more information?

what's in the data file that this label
points to?

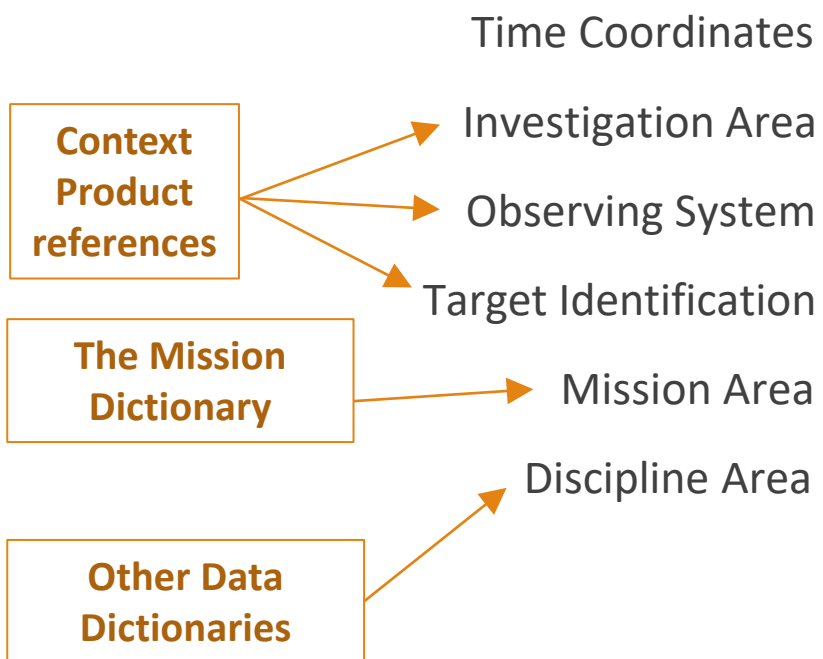
```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<?xml-model ...>
<?xml-model ...>

<Product_Observational ... >
  <Identification_Area> ... </Identification_Area>
  <Observation_Area> ... </Observation_Area>
  <Reference_List> ... </Reference_List>
  <File_Area_Observational> ... </File_Area_Observational>
</Product_Observational>
```

*Observation_Area will probably be the biggest part of your label.
There's a lot going on here.*

How To Design A Label—Part 4, Observation_Area

Subclasses in Observation_Area:



```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<?xml-model ...>
<?xml-model ...>
<Product_Observational ... >
  <Identification_Area> ... </Identification_Area>
  <Observation_Area> ... </Observation_Area>
  <Reference_List> ... </Reference_List>
  <File_Area_Observational> ... </File_Area_Observational>
</Product_Observational>
```

*Before we go into the Observation_Area there will be a slight digression
about Context Products ...*

Digression: Context Products

Context products are descriptions of:

- missions
- spacecraft
- instruments
- targets

that have been registered with PDS.

A data product label includes LIDs of context products to place the data product in context.

They are written by PDS with input from the mission.

Registered context products are here <https://pds.nasa.gov/data/pds4/context-pds4/>

Example context product LIDs for the VIPER mission

Mission	urn:nasa:pds:context:investigation:mission.viper
Spacecraft	urn:nasa:pds:context:instrument_host:spacecraft.viper
Instrument MSOLO	urn:nasa:pds:context:instrument:viper.msolo
Instrument NIRVSS	urn:nasa:pds:context:instrument:viper.nirvss
Instrument NSS	urn:nasa:pds:context:instrument:viper.nss
Instrument TRIDENT	urn:nasa:pds:context:instrument:viper.trident
Instrument VIS	urn:nasa:pds:context:instrument:viper.vis
Instrument Rover?	urn:nasa:pds:context:instrument:viper.rover
Target	urn:nasa:pds:context:target:satellite.earth.moon

Digression: Context Products

Use of a context product LID in a label:

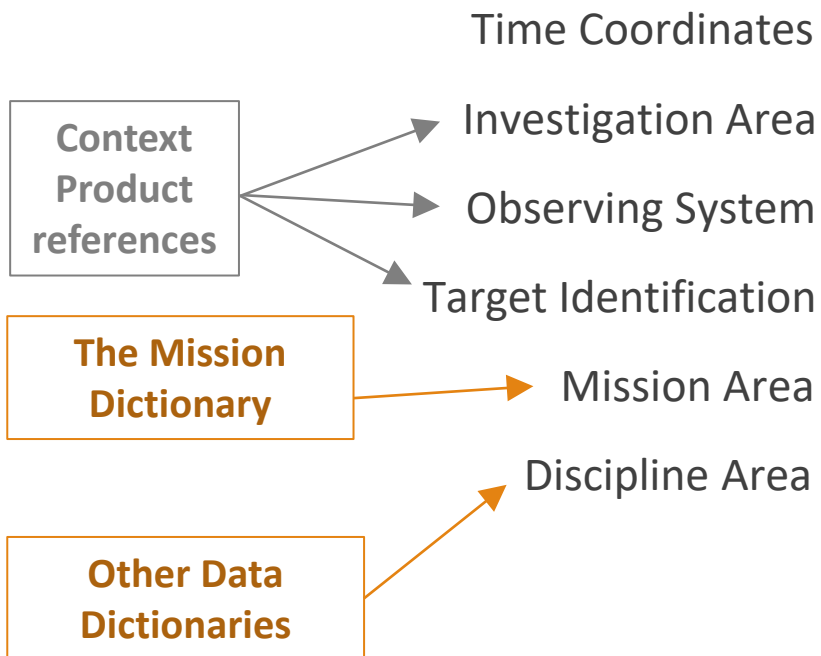
```
...
<Product_Observational ...>
  ...
  <Observation_Area>
    ...
    <Investigation_Area>
      <name>VIPER</name>
      <type>Mission</type>
      <Internal_Reference>
        <lid_reference>
          urn:nasa:pds:context:investigation:mission.viper
        </lid_reference>
        <reference_type>data_to_investigation</reference_type>
      </Internal_Reference>
    </Investigation_Area>
  ...
</Observation_Area>
...
</Product_Observational>
```

The context product itself:

```
...
<Product_Context ...>
  <Identification_Area>
    <logical_identifier>
      urn:nasa:pds:context:investigation:mission.viper
    </logical_identifier>
    <version_id>1.0</version_id>
    <title>Volatiles Investigating Polar Exploration Rover</title>
  ...
</Identification_Area>
<Reference_List>...</Reference_List>
<Investigation>
  <name>VIPER</name>
  <type>Mission</type>
  <start_date>2022-03-31</start_date>
  <stop_date nilReason="unknown" xsi:nil="true"></stop_date>
  <description>
    Insert brief paragraph describing mission
  </description>
</Investigation>
</Product_Context>
```

How To Design A Label—Part 4, Observation_Area

Subclasses in Observation_Area:



```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>  
  
<?xml-model ...>  
<?xml-model ...>  
  
<Product_Observational ... >  
  <Identification_Area> ... </Identification_Area>  
  <Observation_Area> ... </Observation_Area>  
  <Reference_List> ... </Reference_List>  
  <File_Area_Observational> ... </File_Area_Observational>  
</Product_Observational>
```

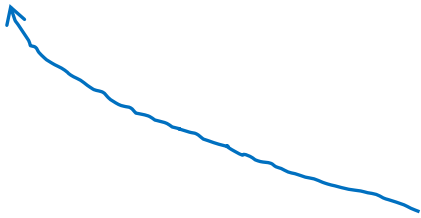
Example of Mission_Area from a Mars 2020 PIXL label

```
<Mission_Area>
  <mars2020:Mars2020_Parameters>
    <mars2020:Observation_Information>
      <mars2020:release_number>0002</mars2020:release_number>
      <mars2020:mission_phase_name>Surface Mission</mars2020:mission_phase_name>
      <mars2020:spacecraft_clock_start>681733043</mars2020:spacecraft_clock_start>
      <mars2020:spacecraft_clock_partition>1</mars2020:spacecraft_clock_partition>
      <mars2020:sol_number>167</mars2020:sol_number>
      <mars2020:start_local_mean_solar_time>Sol-00167M03:39:26.262</mars2020:start_local_mean_solar_time>
      <mars2020:start_local_true_solar_time>03:51:50</mars2020:start_local_true_solar_time>
      <mars2020:start_local_true_solar_time_sol>167</mars2020:start_local_true_solar_time_sol>
      <mars2020:start_solar_longitude unit="deg">82.9305</mars2020:start_solar_longitude>
      <mars2020:active_flight_computer>A</mars2020:active_flight_computer>
    </mars2020:Observation_Information>
    <mars2020:PIXL_Parameters>
      <mars2020:pixl_motion_counter>2139</mars2020:pixl_motion_counter>
      ...
    </mars2020:PIXL_Parameters>
  </mars2020:Mars2020_Parameters>
</Mission_Area>
```

How This Might Work in a VIPER Label

```
<Mission_Area>
  <vipер:VIPER_Parameters>
    <vipер:Observation_Information>
      <vipер:release_number>0002</vipер:release_number>
      <vipер:mission_phase_name>Surface Mission</vipер:mission_phase_name>
      <vipер:spacecraft_clock_start>681733043</vipер:spacecraft_clock_start>
      <vipер:spacecraft_clock_stop>681733051</vipер:spacecraft_clock_stop>
    </vipер:Observation_Information>
    <vipер:NIRVSS_Parameters>
      <vipер:whatsforlunch>BLT with mayo</vipер:whatsforlunch>
      ...
    </vipер:NIRVSS_Parameters>
  </vipер:VIPER_Parameters>
</Mission_Area>
```

This is a silly example just to illustrate that you can create any classes and attributes that make sense to you.



How To Create A Mission Dictionary

PDS will provide a draft of a dictionary definition file and explain how to write definitions.

The file will be on the PDS Dictionary GitHub site for all to access during development. A PDS rep will be the dictionary steward; i.e., the gatekeeper.

At the time of the first data release, the dictionary will be made public.

One way to get started: While designing the Mission_Area of your labels, **use your imagination** and put in whatever terms (attributes and classes) you think need to be there. Just make them up. PDS will go over them with you and see if any them already exist in another PDS dictionary, and if not, we'll define them in the mission dictionary.

Example of Discipline_Area from a Mars 2020 PIXL label

```
...
<Product_Observational ...>
  ...
  <Observation_Area>
    ...
    <Mission_Area>
      ...
      </Mission_Area>
      <Discipline_Area>
        <geom:Geometry> ... </geom:Geometry>
        <msn_surface:Surface_Mission_Information> ... </msn_surface:Surface_Mission_Information>
        <proc:Processing_Information> ... </proc:Processing_Information>
      </Discipline_Area>
    </Observation_Area>
  ...
</Product_Observational>
```

geom, msn_surface, and proc are
PDS Discipline Dictionaries

How This Might Work in a VIPER Label

<Observation_Area>

...

<Mission_Area>

<vipер:VIPER_Parameters>

<vipер:Observation_Information>

...

</vipер:Observation_Information>

<vipер:NIRVSS_Parameters>

...

</vipер:NIRVSS_Parameters>

</vipер:VIPER_Parameters>

</Mission_Area>

<Discipline_Area>

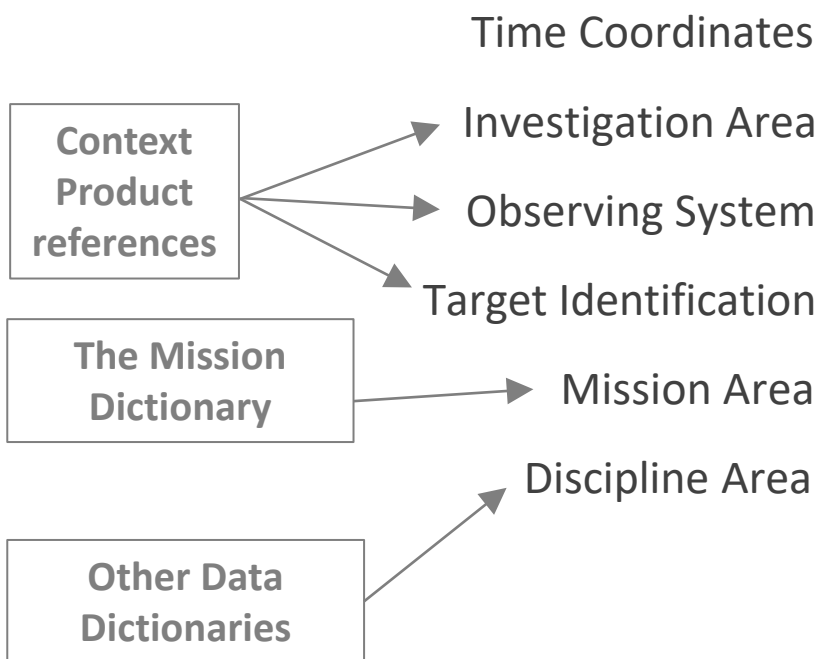
well, it's really hard to tell what goes here without some examples of data products

</Discipline_Area>

</Observation_Area>


How To Design A Label—Part 4, Observation_Area

Subclasses in Observation_Area:



```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<?xml-model ...>
<?xml-model ...>
<Product_Observational ... >
  <Identification_Area> ... </Identification_Area>
  <Observation_Area> ... </Observation_Area>
  <Reference_List> ... </Reference_List>
  <File_Area_Observational> ... </File_Area_Observational>
</Product_Observational>
```

We've now covered Observation_Area and the other three parts of a Product_Observational label, so that's all... except for those mysterious lines at the top.

 Warning: The next slide is for label geeks. Do not be alarmed.

Secrets revealed: what all that stuff at the top of the label is for

```
<?xml-model href="https://pds.nasa.gov/pds4/pds/v1/PDS4_PDS_1G00.sch" ... >
<?xml-model href="https://pds.nasa.gov/pds4/mission/viper/v1/PDS4_VIPER_1G00_1000.sch" ... >
<Product_Observational xmlns="http://pds.nasa.gov/pds4/pds/v1"
  xmlns:pds="http://pds.nasa.gov/pds4/pds/v1"
  xmlns:viper="http://pds.nasa.gov/pds4/mission/viper/v1"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://pds.nasa.gov/pds4/pds/v1
    https://pds.nasa.gov/pds4/pds/v1/PDS4_PDS_1G00.xsd
    http://pds.nasa.gov/pds4/pds/mission/viper/v1 https://pds.nasa.gov/pds4/pds/v1/PDS4_VIPER_1G00_1000.xsd">
  <Mission_Area>
    <viper:VIPER_Parameters>
      <viper:Observation_Information>
        <viper:release_number>0002</viper:release_number>
        <viper:mission_phase_name>Surface Mission</viper:mission_phase_name>
        <viper:spacecraft_clock_start>681733043</viper:spacecraft_clock_start>
        <viper:spacecraft_clock_start>681733051</viper:spacecraft_clock_start>
      </viper:Observation_Information>
      <viper:NIRVSS_Parameters>
        <viper:whatsforlunch>BLT with mayo</viper:whatsforlunch>
      ...
    </viper:NIRVSS_Parameters>
  </viper:VIPER_Parameters>
</Mission_Area>
```

The viper "namespace" is defined here (points to `xmlns:viper="http://pds.nasa.gov/pds4/mission/viper/v1"`)

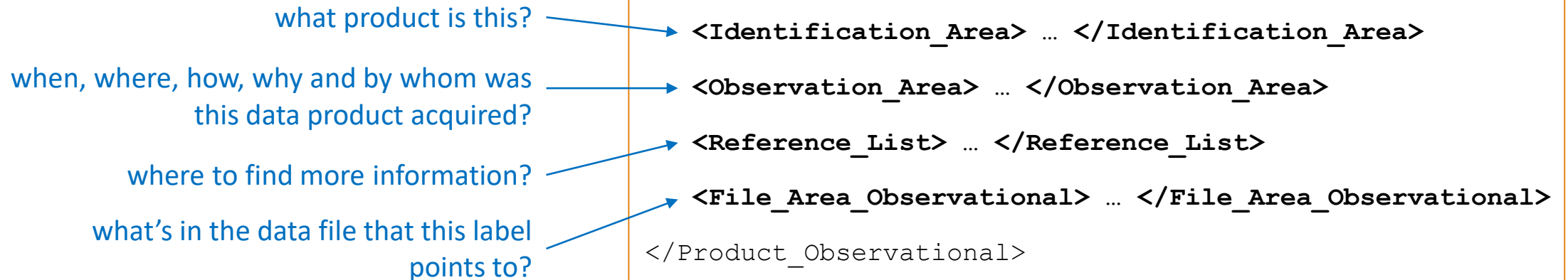
This is the XML Schematron file associated with the dictionary (extension .sch) (points to `https://pds.nasa.gov/pds4/mission/viper/v1/PDS4_VIPER_1G00_1000.sch`)

The namespace is shorthand for this dictionary... (points to `xmlns:viper="http://pds.nasa.gov/pds4/mission/viper/v1"`)

... which is physically located here, in this XML schema file (extension .xsd) (points to `https://pds.nasa.gov/pds4/pds/v1/PDS4_PDS_1G00.xsd` and `https://pds.nasa.gov/pds4/pds/mission/viper/v1 https://pds.nasa.gov/pds4/pds/v1/PDS4_VIPER_1G00_1000.xsd`)

You don't need to understand all this to start drafting a label. Eventually it will become clear.

How To Design A Label—Recap



Outline—where are we?

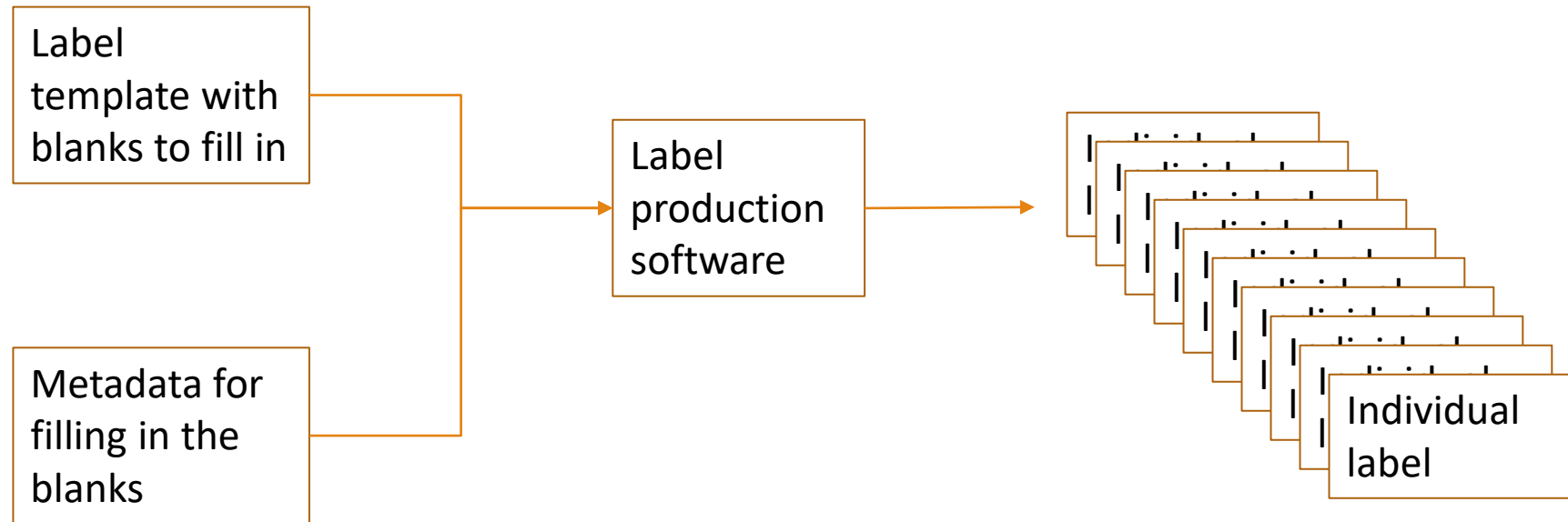
1. What a PDS Label Is For
2. Anatomy of a PDS4 Label
3. Examples Used In This Tutorial
4. How To Design a Label
 - a. Identification_Area
 - b. File_Area_Observational
 - c. Reference_List
 - d. Observation_Area
5. **What To Do With the Label You Have Designed**
 - a. You have a label design, so now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

What To Do With The Label You Have Designed

The following is a brief overview of the next steps in preparing labels for your actual data products. The details are outside the scope of this Label Design Tutorial.

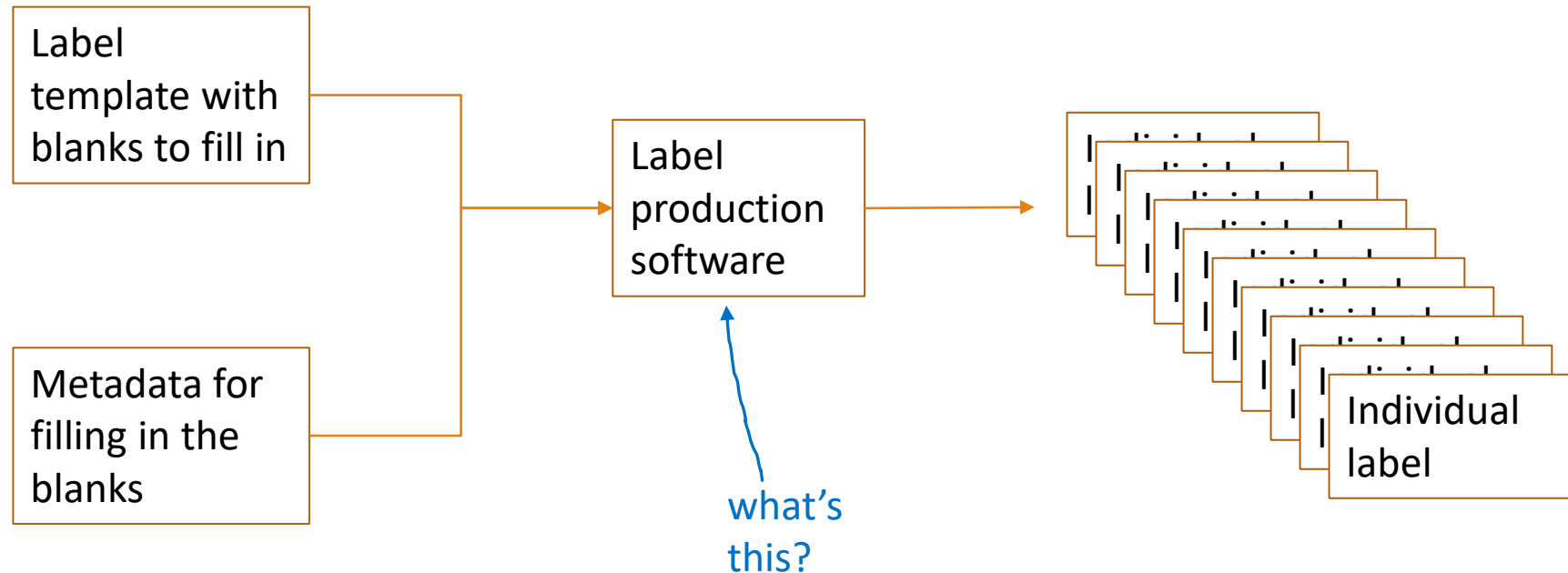
You have a label design, so now what?

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



You have a label design, so now what?

How to go from label design to production of real labels



Label Production Software

Options:

- Your own data processing software
- PDS MI-Label (<https://nasa-pds.github.io/mi-label/>)
- PDS PPI Node Docgen (<http://release.igpp.ucla.edu/igpp/docgen/>)
- PDS Geosciences Node MakeLabels (<https://pds-geosciences.wustl.edu/tools/makelabels.html>)
- ?

Turning a Draft Label into a Label Template

Start with a valid draft label. Identify values that will vary from one label to the next.

```
<Product_Observational ... >
  <Identification_Area>
    <logical_identifier>urn:nasa:pds:viper_nirvss_spc:data_raw:nirvss_sw_nirvss_xt038_sw_throughput_l1_spectra_20220103_194501</logical_identifier>
    <version_id>1.0</version_id>
    <title>VIPER NIRVSS Data Product NIRVSS_SW_NIRVSS_XT038_SW_Throughput_L1_spectra_20220103_194501</title>
    ...
  </Identification_Area>
  <Observation_Area>
    <Time_Coordinates>
      <start_date_time>2021-08-08T22:20:49.776Z</start_date_time>
      <stop_date_time>2021-08-08T22:21:00.001Z</stop_date_time>
    </Time_Coordinates>
    ...
  </Observation_Area>
  ...
  <File_Area_Observational>
    <File>
      <file_name>NIRVSS_SW_NIRVSS_XT038_SW_Throughput_L1_spectra_20220103_194501.csv</file_name>
      <local_identifier>NIRVSS_SW_NIRVSS_XT038_SW_Throughput_L1_spectra_20220103_194501</local_identifier>
      <creation_date_time>2022-01-03T19:45:01Z</creation_date_time>
    </File>
    <Table_Delimited> ... </Table_Delimited>
  </File_Area_Observational>
</Product_Observational>
```

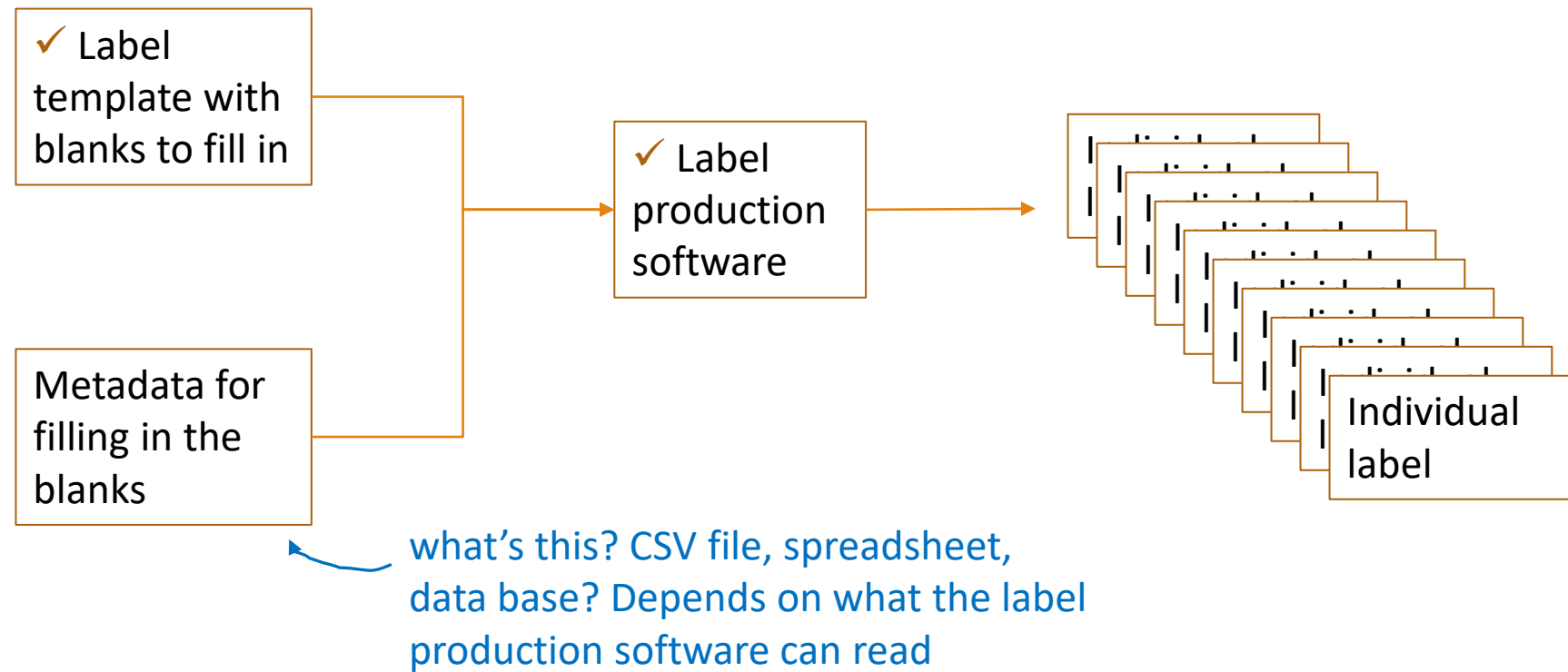
Turning a Draft Label into a Label Template

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



Validating Labels

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

PDS requires that you run 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.

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/