Bridging VESPA and NASA/PDS

Baptiste Cecconi, Pierre Le Sidaner, Cyril Chauvin

- **VESPA:** uses TAP + EPNcore for data discovery.

- **PDS:** NASA planetary science archive

- **Aims:**
  - query for PDS4 resources in VESPA
  - build a VESPA-EPNcore keyword dictionary to allow the use of EPNcore keywords in PDS4

PADC Paris Astronomical Data Centre

europlanet

International Planetary Data Alliance IPDA

NASA PDS v4

- PDS = Planetary Science Archive.
- PDS v4 is an implementation of the Open Archival Information System (OAIS) Reference Model.
- PDS Information model (IM) = data model.
  Archive oriented (preservation, reusability)
- Each PDS entity is a “product” (an object + a label).
- PDS v4 labels are XML files.
- Hierarchy of governance for defining and maintaining dictionaries.

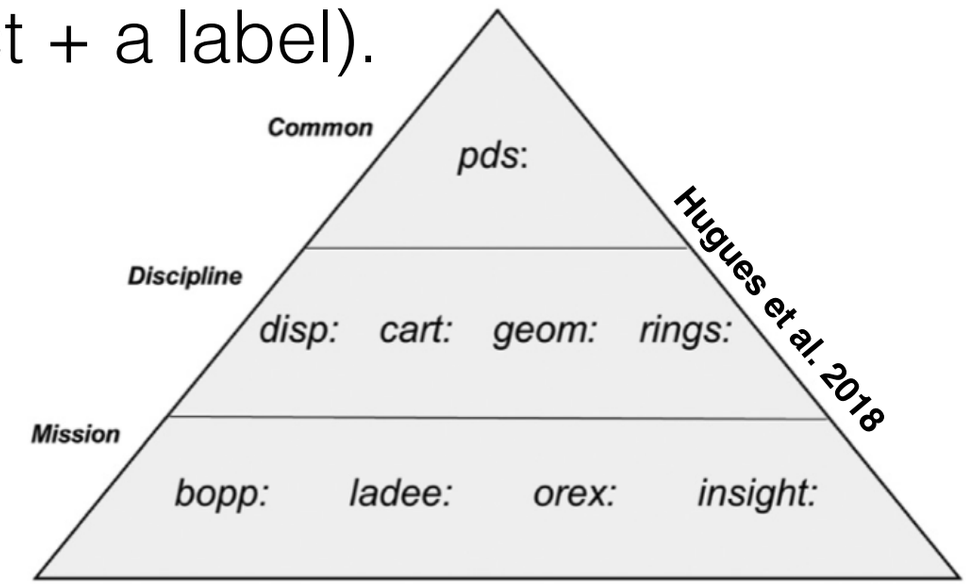


Fig. 1. - Multi-level Governance in the PDS4 Information Model.

PDS Local Data Dictionary

- Local Data Dictionaries (LDD): extension of PDS4 IM
  Possibility to add new concepts into information model.

- LDD = XML schema + Schematron rules.

- LDD is built with LDDtool: input = simple XML file

- See here:
  http://sbndev.astro.umd.edu/wiki/Installing_and_Configuring_LDDTool

VESPA EPNcore data model

- EPNcore = set of keywords for science data products: coverage (ranges, resolutions), observation system, measured quantity, access…
  Data discovery oriented

- Flat data model v2.0: https://voparis-confluence.obspm.fr/display/VES/EPNcore+v2

EPNcore LDD

- This EPN VESPA LDD version propose a nested architecture corresponding to EPN VESPA attributes classification in the EPN VESPA documentation, except the issues mentioned below.

- **Features**:
  - explict c1min, c1max, c2min, c2max, c3min and c3max coordinate names
  - single value specification enabled when range bounds are the same.
  - s_region is decomposed into pieces, making validation possible
  - added _min/_max to subsolar_longitude, subsolar_latitude, subobserver_longitude, subobserver_latitude, ra, dec, radial_distance, altitude_fromshape
  - bib_reference is described (and is validated) either as a bibref, a doi or an url.
  - Detached extension classes:
    + Particle_Spectroscopy_Extension_Parameters,
    + Solar_System_Ojects_Extension_Parameters
    + Experimental_Spectroscopy_Extension_Parameters
    + APIS_Extension_Parameters
  - when *_min and *_max values are equal, the user can select the corresponding *_value attribute instead of the *_min and *_max attributes.

- **Current issues**:
  - c1min, c1max, c2min, c2max, c3min and c3max not defined as the VESPA EPNcore model. Each axis is explicitly defined, with proper names and units.
  - *_value keyword not in VESPA EPNcore model

EPNcore LDD

Ingest_LDD_VESPA.xml – definition of time_min keyword

```
<!-- time axis attributes -->

<DD_Attribute> <!-- time_min [mandatory] -->
    <name>time_min</name>
    <version_id>2.0</version_id>
    <local_identifier>time_min</local_identifier>
    <nillable_flag>true</nillable_flag>
    <submitter_name>Baptiste Cecconi</submitter_name>
    <definition>The time_min parameter provide the start date and time of acquisition in the observer frame.
        Time parameters are provided in UTC and formatted in Julian days (expressed as a double precision
        float).</definition>
    <DD_Value_Domain>
        <enumeration_flag>false</enumeration_flag>
        <value_data_type>ASCII_Real</value_data_type>
        <unit_of_measure_type>Units_of_Time</unit_of_measure_type>
        <specified_unit_id>julian day</specified_unit_id>
    </DD_Value_Domain>
</DD_Attribute>
```

EPNcore LDD

Ingest_LDD_VESPA.xml – definition of time_range class

```
<!-- time_axis classes -->

<DD_Class> <!-- time_range class -->
    <name>time_range</name>
    <version_id>2.0</version_id>
    <local_identifier>time_range</local_identifier>
    <submitter_name>Baptiste Cecconi</submitter_name>
    <definition>The parameters available in this class provides range or value of the date and time of acquisition
        in the observer frame. Use time_min and time_max in case of a range, and time_value for a single event.
        Time parameters are provided in UTC and formatted in Julian days (expressed as a double precision float).
        Although ObsCore uses Modified JD, EPNCore uses standard JD to avoid ambiguity with time origin. With
        double precision floats, the accuracy is on the order of 1 ms, which is considered sufficient to identify
        data of interest (the initial accuracy is preserved in the data itself). The two values min/max permit to
        handle long periods.</definition>

    <DD_Association>
        <identifier_reference>time_value</identifier_reference>
        <reference_type>attribute_of</reference_type>
        <minimum_occurrences>0</minimum_occurrences>
        <maximum_occurrences>1</maximum_occurrences>
    </DD_Association>

    <DD_Association>
        <identifier_reference>time_min</identifier_reference>
        <reference_type>attribute_of</reference_type>
        <minimum_occurrences>0</minimum_occurrences>
        <maximum_occurrences>1</maximum_occurrences>
    </DD_Association>

    <DD_Association>
        <identifier_reference>time_max</identifier_reference>
        <reference_type>attribute_of</reference_type>
        <minimum_occurrences>0</minimum_occurrences>
        <maximum_occurrences>1</maximum_occurrences>
    </DD_Association>
</DD_Class>
```

EPNcore LDD

Ingest_LDD_VESPA.xml – definition of schematron rules

```
<!-- Schematron rules -->

<DD_Rule> <!-- check_time_range_value -->
    <local_identifier>check_time_range_value</local_identifier>
    <rule_context>epn:VESPA_EPNcore_Parameters/epn:time_axis/epn:time_range</rule_context>
    <DD_Rule_Statement>
        <rule_type>Assert</rule_type>
        <rule_test>boolean(epn:time_value) != boolean(epn:time_min or epn:time_max)</rule_test>
        <rule_message>Use either time_value or time_min and time_max</rule_message>
    </DD_Rule_Statement>
</DD_Rule>
<DD_Rule> <!-- check_time_bounds -->
    <local_identifier>check_time_bounds</local_identifier>
    <rule_context>epn:VESPA_EPNcore_Parameters/epn:time_axis/epn:time_range</rule_context>
    <DD_Rule_Statement>
        <rule_type>Assert</rule_type>
        <rule_test>not(epn:time_min) or not(epn:time_max) or epn:time_min &amp;lt;= epn:time_max</rule_test>
        <rule_message>time_min must be lower or equal than time_max</rule_message>
    </DD_Rule_Statement>
</DD_Rule>
```

EPNcore LDD

Ingest_LDD_VESPA.xml – definition of spatial axis keywords

```
<!-- spatial axes attributes -->

<DD_Attribute> <!-- spatial_frame_type [mandatory] --> [60 lines]

<DD_Attribute> <!-- ra_min (c1 in celestial frame) [mandatory] -->
    <name>ra_min</name>
    <version_id>2.0</version_id>
    <local_identifier>ra_min</local_identifier>
    <nillable_flag>true</nillable_flag>
    <submitter_name>Baptiste Cecconi</submitter_name>
    <definition>The ra_min (c1min in celestial frame) parameter provides the lowest bound of the Right Ascension
        celestial coordinate.
    </definition>
    <DD_Value_Domain>
        <enumeration_flag>false</enumeration_flag>
        <value_data_type>ASCII_Real</value_data_type>
        <minimum_value>0</minimum_value>
        <maximum_value>360</maximum_value>
        <unit_of_measure_type>Units_of_Angle</unit_of_measure_type>
        <specified_unit_id>deg</specified_unit_id>
    </DD_Value_Domain>
</DD_Attribute>
```

EPNcore LDD

Ingest_LDD_VESPA.xml – definition of schematron rules

```
<DD_Rule> <!-- check_spatial_frame_type_and_class -->
    <local_identifier>check_spatial_frame_type_and_class</local_identifier>
    <rule_context>epn:VESPA_EPNcore_Parameters/epn:spatial_axes</rule_context>
    <DD_Rule_Statement>
        <rule_type>Assert</rule_type>
        <rule_test>epn:spatial_frame_type = 'none' or
            (epn:spatial_frame_type = 'celestial' and epn:celestial_frame_axes) or
            (epn:spatial_frame_type = 'body' and epn:body_frame_axes) or
            (epn:spatial_frame_type = 'cartesian' and epn:cartesian_frame_axes) or
            (epn:spatial_frame_type = 'spherical' and epn:spherical_frame_axes) or
            (epn:spatial_frame_type = 'sylindrical' and epn:cylindrical_frame_axes)</rule_test>
        <rule_message>The *_frame_axes element corresponding to the spatial_frame_type must be present</rule_message>
    </DD_Rule_Statement>
</DD_Rule>
```

EPNcore LDD

Example_label.xml – section of file, with spatial keywords

```
<spatial_axes>
    <spatial_frame_type>cartesian</spatial_frame_type>
    <spatial_coordinate_description>Jupiter_III</spatial_coordinate_description>
    <spatial_origin>Jupiter</spatial_origin>
    <cartesian_frame_axes>
        <x_min unit="m">10</x_min>
        <x_max unit="m">100</x_max>
        <y_min unit="m">-5</y_min>
        <y_max unit="m">5</y_max>
        <z_min unit="m">100</z_min>
        <z_max unit="m">10000</z_max>
    </cartesian_frame_axes>
    <s_region>
        <s_region_type>polygon</s_region_type>
        <s_region_polygon_point>
            <s_region_lon unit="deg">10</s_region_lon>
            <s_region_lat unit="deg">-5</s_region_lat>
        </s_region_polygon_point>
        <s_region_polygon_point>
            <s_region_lon unit="deg">10</s_region_lon>
            <s_region_lat unit="deg">5</s_region_lat>
        </s_region_polygon_point>
```

EPNcore LDD

Example_label.xml – begin of file, with time_axis keywords (ranges and and value)

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model href="Ingest_LDD_VESPA_EPN_1x8000.sch" type="application/xml" schematypens="http://purl.oclc.org/dsdl/schematron"?>
<VESPA_EPNcore_Parameters xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xmlns="http://pds.nasa.gov/pds4/epn/v1"
    xsi:schemaLocation="http://pds.nasa.gov/pds4/epn/v1 Ingest_LDD_VESPA_EPN_1x8000.xsd">
    <granule_uid>o4bda5kaq_1D_pdf</granule_uid>
    <granule_gid>1D_spectra</granule_gid>
    <obs_id>o4bda5kaq</obs_id>
    <service_title>APIS</service_title>
    <dataproduct_type>sp</dataproduct_type>
    <measurement_type>phot.flux.density;em.UV.100-200nm</measurement_type>
    <processing_level>5</processing_level>
    <target_class>planet</target_class>
    <target_name>Saturn</target_name>
    <target_region>magnetosphere</target_region>
    <feature_name>planetary aurorae</feature_name>
    <instrument_host_name>hst</instrument_host_name>
    <instrument_name>STIS</instrument_name>
    <filter>52X0.5</filter>
    <time_axis>
        <time_range>
            <time_min unit="julian day">2450787.7134216</time_min>
            <time_max unit="julian day">2450787.72499787</time_max>
        </time_range>
        <time_exp>
            <time_exp_value unit="s">1000.1899</time_exp_value>
        </time_exp>
    </time_axis>
```

Real use case
Cassini/RPWS/HFR/QTN collection

- PDS4 Archive bundle preparation for Cassini/RPWS/HFR/QTN

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model href="https://pds.nasa.gov/pds4/pds/v
    schematypens="http://purl.oclc.org/dsdl/schema
    type="application/xml"?><Product_Observational
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-in
    xsi:schemaLocation="http://pds.nasa.gov/pds4/p
    http://pds.nasa.gov/pds4/epn/v1 Ingest_LDD_VES
    <Identification_Area>
        <logical_identifier>urn:nasa:pds:co-rpws-s
        <version_id>1.0</version_id>
        <title>Quasi Thermal Noise pectroscopy on
        <information_model_version>1.10.1.0</infor
        <product_class>Product_Observational</proc
        <Citation_Information>
            <publication_year>2018</publication_ye
            <description>This product contains RPW
                containing a time series of therma
                are generated only when the QTN ar
        </Citation_Information>
        <Modification_History>
            <Modification_Detail>
                <modification_date>2018-11-06</mod
                <version_id>1.0</version_id>
                <description>HFR-QTN Initial Relea
            </Modification_Detail>
        </Modification_History>
    </Identification_Area>
    <Observation_Area>
        <Time_Coordinates>
            <start_date_time>2004-06-30T16:42:09.8
            <stop_date_time>2004-07-01T09:51:13.41
        </Time_Coordinates>
        <Investigation_Area>
            <name>Cassini-Huygens</name>
            <type>Mission</type>
            <Internal_Reference>
                <lid_reference>urn:nasa:pds:contex
                <reference_type>data_to_investigat
            </Internal_Reference>
        </Investigation_Area>
        <Observing_System>
            <Observing_System_Component>
                <name>Cassini Orbiter</name>
                <type>Spacecraft</type>
```

```
        <Discipline_Area>
            <epn:VESPA_EPNcore_Parameters>
                <epn:granule_uid>urn:nasa:pds:co-rpws-saturn:hfr-qtn-data:200406301642-200407010951-cdf</epn:granule_uid>
                <epn:granule_gid>co-rpws-hfr-qtn</epn:granule_gid>
                <epn:obs_id>co-rpws</epn:obs_id>
                <epn:service_title/>
                <epn:dataproduct_type>ts</epn:dataproduct_type>
                <epn:measurement_type>phys.density;phys.electron#phys.temperature;phys.electron</epn:measurement_type>
                <epn:measurement_type/>
                <epn:processing_level>5</epn:processing_level>
                <epn:target_class>planet</epn:target_class>
                <epn:target_name>Saturn</epn:target_name>
                <epn:time_axis>
                    <epn:time_sampling_step>
                        <epn:time_sampling_step_min unit="s">12.619967</epn:time_sampling_step_min>
                        <epn:time_sampling_step_max unit="s">10127.939996</epn:time_sampling_step_max>
                    </epn:time_sampling_step>
                </epn:time_axis>
                <epn:spatial_axes>
                    <epn:spatial_frame_type>spherical</epn:spatial_frame_type>
                    <epn:spherical_frame_axes>
                        <epn:r_min unit="km"></epn:r_min>
                        <epn:r_max unit="km"></epn:r_max>
                        <epn:colat_min unit="deg">85.9672742073</epn:colat_min>
                        <epn:colat_max unit="deg">101.924017848</epn:colat_max>
                    </epn:spherical_frame_axes>
                    <epn:spatial_coordinate_description>SSQ - Saturn Solar Equatorial</epn:spatial_coordinate_description>
                </epn:spatial_axes>
                <epn:other_location>
                    <epn:local_time>
                        <epn:local_time_min unit="hr">0.00509811014416</epn:local_time_min>
                        <epn:local_time_max unit="hr">23.9999247824</epn:local_time_max>
                    </epn:local_time>
                </epn:other_location>
            </epn:VESPA_EPNcore_Parameters>
        </Discipline_Area>
```

Starting the Mapping

Using **Terminology_Entry_SKOS** to map between dictionaries

```
7
8     <!-- VESPA_EPNcore Attributes -->
9
10    <!-- [[ epn:granule_uid ]] -->
11
12    <!-- epn:granule_uid is a _exactMatch_ to pds:Product_Components/Identification_Area/logical_identifier -->
13    <Terminological_Entry_SKOS>
14        <identifier>epn.granule_uid</identifier>
15        <namespace_id>epn</namespace_id>
16        <steward_id>epn</steward_id>
17        <title>granule_uid</title>
18        <description>Unique granule ID in collection: logical_identifier of a Product</description>
19        <referenced_identifier>logical_identifier</referenced_identifier>
20        <skos_relation_name>exactMatch</skos_relation_name>
21        <model_object_id>0001_NASA_PDS_1.pds:Product_Components/Identification_Area/logical_identifier</model_object_id>
22        <model_object_type>attribute</model_object_type>
23    </Terminological_Entry_SKOS>
24
```

Starting the Mapping

```
80
81      <Terminological_Entry_SKOS>
82          <identifier>epn.target_class.planet</identifier>
83          <namespace_id>epn</namespace_id>
84          <steward_id>epn</steward_id>
85          <title>planet</title>
86          <description>Class of a target that is a Planet</description>
87          <referenced_identifier>Planet</referenced_identifier>
88          <skos_relation_name>exactMatch</skos_relation_name>
89          <model_object_id>0001_NASA_PDS_1.pds:Target_Identification/type:Planet</mo
90          <model_object_type>attribute</model_object_type>
91      </Terminological_Entry_SKOS>
92
93      <Terminological_Entry_SKOS>
94          <identifier>epn.target_class.asteriod</identifier>
95          <namespace_id>epn</namespace_id>
96          <steward_id>epn</steward_id>
97          <title>asteriod</title>
98          <description>Class of a target that is an Asteroid</description>
99          <referenced_identifier>Asteriod</referenced_identifier>
100         <skos_relation_name>broadMatch</skos_relation_name>
101         <model_object_id>0001_NASA_PDS_1.pds:Target_Identification/type:Asteroid</
102         <model_object_type>attribute</model_object_type>
103     </Terminological_Entry_SKOS>
104
105     <Terminological_Entry_SKOS>
106         <identifier>epn.target_class.asteriod</identifier>
107         <namespace_id>epn</namespace_id>
108         <steward_id>epn</steward_id>
109         <title>asteriod</title>
110         <description>Class of a target that is an Asteroid</description>
111         <referenced_identifier>Trans-Neptunian Object</referenced_identifier>
112         <skos_relation_name>broadMatch</skos_relation_name>
113         <model_object_id>0001_NASA_PDS_1.pds:Target_Identification/type:Trans-Nept
114         <model_object_type>attribute</model_object_type>
115     </Terminological_Entry_SKOS>
```

exactMatch:
epn.target_class.planet
to
pds:Target_Indentification/type:Planet

broadMatch:
epn.target_class.asteroid
to
pds:Target_Indentification/type:Asteroid
and
pds:Target_Indentification/type:Trans-Neptunian Object

exactMatch (or closeMatch) ?
epn.target_class.star
to
pds:Target_Indentification/type:Star

narrowMatch:
epn.target_class.star
to
pds:Target_Indentification/type:Sun

Starting the Mapping

```
108     <!-- [[ epn:target_class ]] -->
109
110     <!-- epn:target_class is a _closeMatch_ to pds:Target_Identification/type -->
111     <Terminological_Entry_SKOS> [10 lines]
122
123     <!-- epn:target_class.asteroid is a _broadMatch_ to pds:Target_Identification/type:Asteroid -->
124     <Terminological_Entry_SKOS> [10 lines]
135
136     <!-- epn:target_class.asteroid is a _broadMatch_ to pds:Target_Identification/type:Trans-Neptunian Object -->
137     <Terminological_Entry_SKOS>   [10 lines]
148
149     <!-- epn:target_class.comet is an _exactMatch_ to pds:Target_Identification/type:Comet -->
150     <Terminological_Entry_SKOS> [10 lines]
161
162     <!-- epn:target_class.dwarf_planet is a _broadMatch_ to pds:Target_Identification/type:Asteroid -->
163     <Terminological_Entry_SKOS> [10 lines]
174
175     <!-- epn:target_class.dwarf_planet is a _broadMatch_ to pds:Target_Identification/type:Centaur -->
176     <Terminological_Entry_SKOS> [10 lines]
187
188     <!-- epn:target_class.exoplanet is a _closeMatch_ to pds:Target_Identification/type:Planetary System -->
189     <Terminological_Entry_SKOS> [10 lines]
200
201     <!-- epn:target_class.interplanetary_medium is a _closeMatch_ to pds:Target_Identification/type:Dust -->
202     <Terminological_Entry_SKOS> [10 lines]
213
214     <!-- epn:target_class.interplanetary_medium is a _closeMatch_ to pds:Target_Identification/type:Plasma Stream -->
215     <Terminological_Entry_SKOS> [10 lines]
226
227     <!-- epn:target_class.interplanetary_medium is a _closeMatch_ to pds:Target_Identification/type:Plasma Cloud -->
228     <Terminological_Entry_SKOS> [10 lines]
239
240     <!-- epn:target_class.planet is an _exactMatch_ to pds:Target_Identification/type:Planet -->
241     <Terminological_Entry_SKOS> [10 lines]
252
253     <!-- epn:target_class.planet is an _closeMatch_ to pds:Target_Identification/type:Planetary System -->
254     <Terminological_Entry_SKOS> [10 lines]
265
```

Next steps

- Select option to implement (personal preference = option 5)

- Check on real cases.
  E.g.: Cassini/RPWS/QTN Higher level data collection
  => goal: use VESPA_EPNcore_LDD on this data collection and prototype query of EPNcore keywords using PDS4 search interface.

- Build PDS4 “mapping” product between EPNcore and PDS IM (and other LDDs)
  => include SKOS into Ingest_LDD?

- Leads to improved specification of EPNcore model?

Using PDS Search Interface (current one)

- Prototype PDS4 search interface in VESPA portal
  => What can be found in NASA PDS data.

- We use parameters that can be mapped, convert what can be

- Transform the TAP/ADQL query into Lucene/Solr like query

- Visualize result in a separate web page

Using PDS Search Interface

Rough correspondance between VESPA dictionary and PDS:

target_name= X -> target:X
target_class:
List of possible values ->https://pds.jpl.nasa.gov/tools/dd-search/
* asteroid -> (target-type:ASTEROID OR target-type:TRANS-NEPTUNIAN OBJ)
* comet -> target-type:COMET
* dwarf_planet -> target-type:ASTEROID
* exoplanet -> (target-type:PLANETARY_SYSTEM OR target-type:PLANETARY SYSTEM)
* interplanetary_medium -> target-type:DUST
* planet -> target-type:PLANET
* ring -> target-type:RING
* sample -> (target-type:TERRESTRIAL SAMPLE OR target-type:METEORITE)
* satellite -> target-type:SATELLITE
* sky -> (target-type:GALAXY OR target-type:GLOBULAR CLUSTER
OR target-type:DUST OR target-type:NEBULA OR target-type:
OPEN CLUSTER OR target-type:PLANETARY NEBULA OR
target-type:REFERENCE OR target-type:STAR OR target-type:STAR CLUSTER)
* star -> if target_name =='sun':
target-type:SUN
else:
(target-type:STAR OR target-type:STAR CLUSTER)
* calibration -> target-type:CALIBRATION

instrument_name = X -> (instrument:X OR observing-system:X)
instrument_host_name = X -> (instrument-host:X OR investigation:X)
file_name = X -> title:X

time_min = X -> stop-time:[X TO MAX_ISO_DATE]
time_max = X -> start-time:[MIN_ISO_DATE TO X]

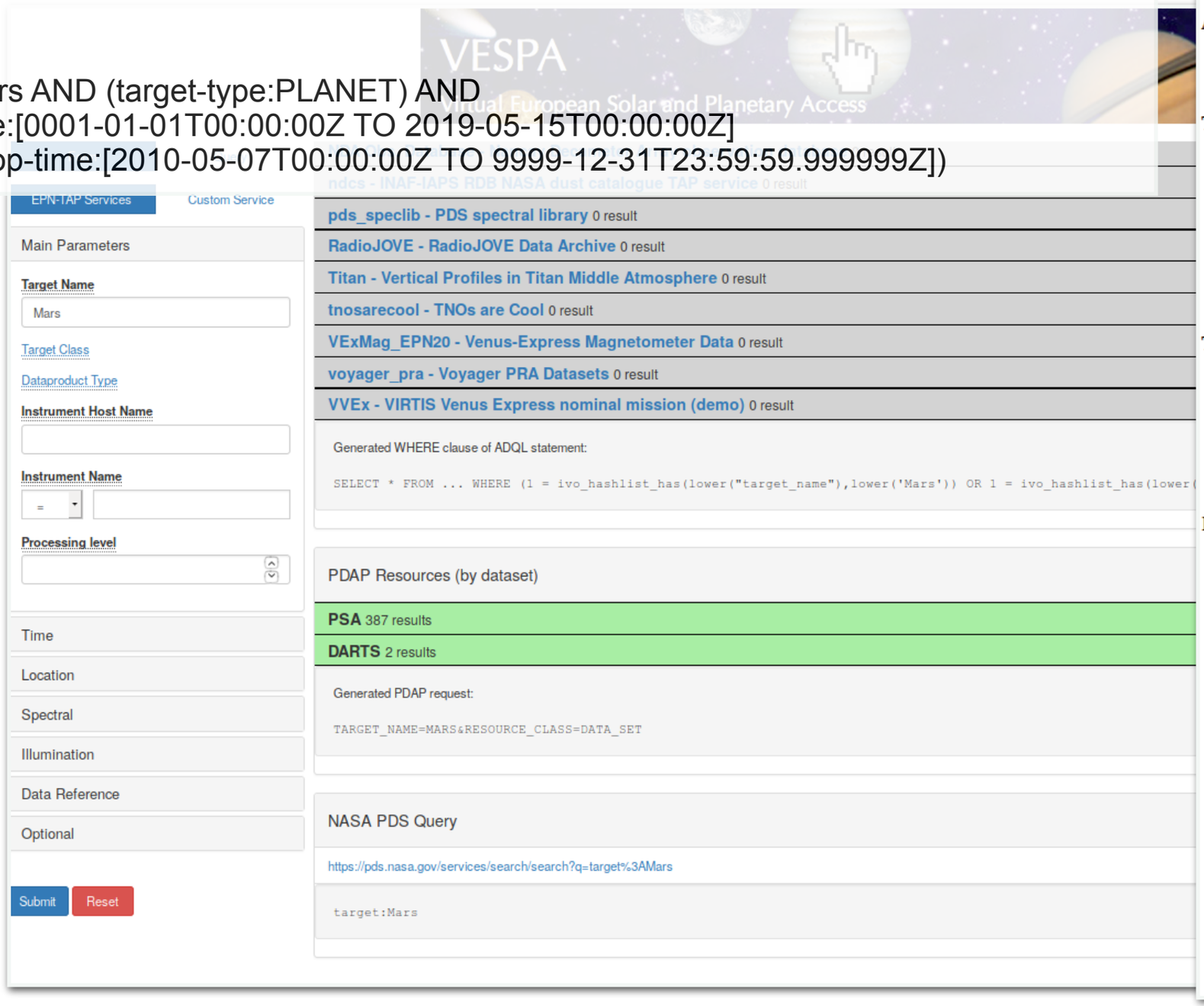
Using PDS Search Interface

Simple Translation Example:

target_name = Mars, target_class=Planet,
time_min = 2010-05-07T00:00:00.000, time_max=2019-05-15T00:00:00.000

Become

target:mars AND (target-type:PLANET) AND
(start-time:[0001-01-01T00:00:00Z TO 2019-05-15T00:00:00Z]
AND stop-time:[2010-05-07T00:00:00Z TO 9999-12-31T23:59:59.999999Z])



EPNcore into PDS LDD

option	name	features	issues
Option 1	original flat	simple EPNcore translation	coordinates, ranges, units, validation
Option 2	original nested	grouped keywords fix: ranges, units	coordinates, validation
~~Option 3~~	~~multidictionary~~	~~several dictionaries to fix multiple definitions of same keyword name~~	~~not implemented~~
Option 4	detailed coords	explicit coordinate names	coordinates definitions differ from EPNCore model
Option 5	detailed coords single values	explicit coordinates names explicit single values/ranges	coordinates definitions differ from EPNCore model

EPNcore LDD option 1 Original Flat schema

- This EPN VESPA LDD version propose a flat architecture corresponding to EPN VESPA attributes.

- **Features**:
  - all EPN VESPA attributes at same level in XML label

- **Current issues**:
  - c1min, c1max, c2min, c2max, c3min and c3max can't have @unit. This will be problematic for mapping with other LDDs, and make validation more complex.
  - s_region is provided is an STC string. No validation process.
  - bib_reference should be split in various subtypes for validation
  - particle_spectral_*_min and particle_spectral_*_max units can't be described (no @unit)
  - access_format missing ASCII_Mime_Type value_data_type
  - attributes missing _min/_max split: subsolar_longitude, subsolar_latitude, subobserver_longitude, subobserver_latitude, ra, dec, radial_distance,altitude_fromshape, mass, sideral_rotation_period, mean_radius, equatorial_radius, polar_radius

EPNcore LDD option 1
Original Flat schema

Ingest_LDD_VESPA.xml – definition of time_min keyword

```
<!-- time axis attributes -->

<DD_Attribute> <!-- time_min [mandatory] -->
    <name>time_min</name>
    <version_id>2.0</version_id>
    <local_identifier>time_min</local_identifier>
    <nillable_flag>true</nillable_flag>
    <submitter_name>Baptiste Cecconi</submitter_name>
    <definition>The time_min parameter provide the start date and time of acquisition in the observer frame.
        Time parameters are provided in UTC and formatted in Julian days (expressed as a double precision
        float).</definition>
    <DD_Value_Domain>
        <enumeration_flag>false</enumeration_flag>
        <value_data_type>ASCII_Real</value_data_type>
        <unit_of_measure_type>Units_of_Time</unit_of_measure_type>
        <specified_unit_id>julian day</specified_unit_id>
    </DD_Value_Domain>
</DD_Attribute>
```

EPNcore LDD option 1
Original Flat schema

Ingest_LDD_VESPA.xml – definition of schematron rules

```
<!-- Schematron rules -->

<DD_Rule> <!-- check_time_range_bounds -->
    <local_identifier>check_time_range_bounds</local_identifier>
    <rule_context>epn:VESPA_EPNcore_Parameters</rule_context>
    <DD_Rule_Statement>
        <rule_type>Assert</rule_type>
        <rule_test>not(epn:time_min) or not(epn:time_max) or epn:time_min &amp;lt;= epn:time_max</rule_test>
        <rule_message>time_min must be lower or equal than time_max</rule_message>
    </DD_Rule_Statement>
</DD_Rule>
<DD_Rule> <!-- check_time_sampling_step_bounds --> [9 lines]
<DD_Rule> <!-- check_time_exp_bounds --> [9 lines]
```

EPNcore LDD option 1
Original Flat schema

Example_label.xml – begin of fil, with several keywords

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model href="Ingest_LDD_VESPA_EPN_1x8000.sch" type="application/xml" schematypens="http://purl.oclc.org/dsdl/schematron"?>
<VESPA_EPNcore_Parameters xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xmlns="http://pds.nasa.gov/pds4/epn/v1"
    xsi:schemaLocation="http://pds.nasa.gov/pds4/epn/v1 Ingest_LDD_VESPA_EPN_1x8000.xsd">
    <granule_uid>test</granule_uid>
    <granule_gid>test</granule_gid>
    <obs_id>test</obs_id>
    <dataproduct_type>ts</dataproduct_type>
    <measurement_type>phys.magField</measurement_type>
    <processing_level>2</processing_level>
    <target_name>Jupiter</target_name>
    <target_class>planet</target_class>
    <time_min unit="julian day">2455000.5</time_min>
    <time_max unit="julian day">2455001.5</time_max>
    <time_sampling_step_min unit="s">1</time_sampling_step_min>
    <time_sampling_step_max unit="s">10</time_sampling_step_max>
    <time_exp_min unit="s">5</time_exp_min>
    <time_exp_max unit="s">5</time_exp_max>
```

EPNcore LDD option 2
Original nested schema

- This EPN VESPA LDD version propose a nested architecture corresponding to EPN VESPA attributes classification in the EPN VESPA documentation

- **Features**:
  - nested EPN VESPA attributes in XML label
  - s_region is decomposed into pieces, making validation possible
  - added _min/_max to subsolar_longitude, subsolar_latitude, subobserver_longitude, subobserver_latitude, ra, dec, radial_distance, altitude_fromshape
  - bib_reference is describes either as a bibref, a doi or an url.

- **Current issues**:
  - c1min, c1max, c2min, c2max, c3min and c3max names are defined several times (with different local_identifier).
  This is confuses the validation engine and false errors on units are reported.

EPNcore LDD option 2
Original nested schema

Ingest_LDD_VESPA.xml – definition of time_range class

```
<!-- time_axis classes -->

<DD_Class> <!-- time_range class -->
    <name>time_range</name>
    <version_id>2.0</version_id>
    <local_identifier>time_range</local_identifier>
    <submitter_name>Baptiste Cecconi</submitter_name>
    <definition>The time_min and time_max parameter provide the start date and time of acquisition in the
        observer frame. Time parameters are provided in UTC and formatted in Julian days (expressed as a
        double precision float). Although ObsCore uses Modified JD, EPNCore uses standard JD to avoid ambiguity
        with time origin. With double precision floats, the accuracy is on the order of 1 ms, which is considered
        sufficient to identify data of interest (the initial accuracy is preserved in the data itself).
        The two values min/max permit to handle long periods. Whenever acquisition time is a scalar (rather than
        an interval), both time_min and time_max must contain the same value in the table. There is no limiting
        value to this parameter</definition>

    <DD_Association>
        <identifier_reference>time_min</identifier_reference>
        <reference_type>attribute_of</reference_type>
        <minimum_occurrences>0</minimum_occurrences>
        <maximum_occurrences>1</maximum_occurrences>
    </DD_Association>

    <DD_Association>
        <identifier_reference>time_max</identifier_reference>
        <reference_type>attribute_of</reference_type>
        <minimum_occurrences>0</minimum_occurrences>
        <maximum_occurrences>1</maximum_occurrences>
    </DD_Association>
</DD_Class>
```

EPNcore LDD option 2
Original nested schema

Ingest_LDD_VESPA.xml – definition of schematron rules

```
<!-- Schematron rules -->

<DD_Rule> <!-- check_time_range_bounds -->
    <local_identifier>check_time_range_bounds</local_identifier>
    <rule_context>epn:VESPA_EPNcore_Parameters</rule_context>
    <DD_Rule_Statement>
        <rule_type>Assert</rule_type>
        <rule_test>not(epn:time_min) or not(epn:time_max) or epn:time_min &amp;lt;= epn:time_max</rule_test>
        <rule_message>time_min must be lower or equal than time_max</rule_message>
</DD_Rule_Statement>
```

```
<!-- Schematron rules -->

<DD_Rule> <!-- check_time_range_bounds -->
    <local_identifier>check_time_frame_bounds</local_identifier>
    <rule_context>epn:VESPA_EPNcore_Parameters/epn:time_axis/epn:time_range</rule_context>
    <DD_Rule_Statement>
        <rule_type>Assert</rule_type>
        <rule_test>not(epn:time_min) or not(epn:time_max) or epn:time_min &amp;lt;= epn:time_max</rule_test>
        <rule_message>time_min must be lower or equal than time_max</rule_message>
    </DD_Rule_Statement>
</DD_Rule>
<DD_Rule> <!-- check_time_sampling_step_bounds --> [9 lines]
<DD_Rule> <!-- check_time_exp_bounds --> [9 lines]
```

EPNcore LDD option 2
Original nested schema

Example_label.xml – begin of file, with several keywords

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model href="Ingest_LDD_VESPA_EPN_1x8000.sch" type="application/xml" schematypens="http://purl.oclc.org/dsdl/schematron"?>
<VESPA_EPNcore_Parameters xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xmlns="http://pds.nasa.gov/pds4/epn/v1" xsi:schemaLocation="http://pds.nasa.gov/pds4/epn/v1 Ingest_LDD_VESPA_EPN_1x8000.xsd">
    <granule>
        <granule_uid>test</granule_uid>
        <granule_gid>test</granule_gid>
        <obs_id>test</obs_id>
</granule>
<data_description>
    <dataproduct_type>ts</dataproduct_type>
    <measurement_type>phys.magField</measurement_type>
    <processing_level>2</processing_level>
</data_description>
<target>
    <target_class>planet</target_class>
    <target_name>Jupiter</target_name>
</target>
<observation_system>
    <instrument_host_name>Cassini</instrument_host_name>
    <instrument_name>Mag</instrument_name>
</observation_system>
<time_axis>
    <time_range>
        <time_min unit="julian day">2455000.5</time_min>
        <time_max unit="julian day">2455001.5</time_max>
    </time_range>
    <time_sampling_step>
        <time_sampling_step_min unit="s">10</time_sampling_step_min>
        <time_sampling_step_max unit="s">10</time_sampling_step_max>
    </time_sampling_step>
    <time_exp>
        <time_exp_min unit="s">5</time_exp_min>
        <time_exp_max unit="s">5</time_exp_max>
    </time_exp>
</time_axis>
```

EPNcore LDD option 4
Detailed coords schema

- This EPN VESPA LDD version propose a nested architecture corresponding to EPN VESPA attributes classification in the EPN VESPA documentation, except for the c1min, c1max, c2min, c2max, c3min and c3max elements, that are replaced by explicit attribute names.

- **Features**:
  - nested EPN VESPA attributes in XML label
  - explict c1min, c1max, c2min, c2max, c3min and c3max coordinate names
  - s_region is decomposed into pieces, making validation possible
  - added _min/_max to subsolar_longitude, subsolar_latitude, subobserver_longitude, subobserver_latitude, ra, dec, radial_distance, altitude_fromshape
  - bib_reference is described (and is validated) either as a bibref, a doi or an url.
  - Detached extension classes:
    + Particle_Spectroscopy_Extension_Parameters,
    + Solar_System_Ojects_Extension_Parameters
    + Experimental_Spectroscopy_Extension_Parameters
    + APIS_Extension_Parameters

- **Current issues**:
  - c1min, c1max, c2min, c2max, c3min and c3max not defined as the VESPA EPNcore model

EPNcore LDD option 4
Detailed coords schema

Ingest_LDD_VESPA.xml – definition of spatial axis keywords

```
<!-- spatial axes attributes -->

<DD_Attribute> <!-- spatial_frame_type [mandatory] --> [60 lines]

<DD_Attribute> <!-- ra_min (c1 in celestial frame) [mandatory] -->
    <name>ra_min</name>
    <version_id>2.0</version_id>
    <local_identifier>ra_min</local_identifier>
    <nillable_flag>true</nillable_flag>
    <submitter_name>Baptiste Cecconi</submitter_name>
    <definition>The ra_min (c1min in celestial frame) parameter provides the lowest bound of the Right Ascension
        celestial coordinate.
    </definition>
    <DD_Value_Domain>
        <enumeration_flag>false</enumeration_flag>
        <value_data_type>ASCII_Real</value_data_type>
        <minimum_value>0</minimum_value>
        <maximum_value>360</maximum_value>
        <unit_of_measure_type>Units_of_Angle</unit_of_measure_type>
        <specified_unit_id>deg</specified_unit_id>
    </DD_Value_Domain>
</DD_Attribute>
```

EPNcore LDD option 4
Detailed coords schema

Ingest_LDD_VESPA.xml – definition of schematron rules

```
<DD_Rule> <!-- check_spatial_frame_type_and_class -->
    <local_identifier>check_spatial_frame_type_and_class</local_identifier>
    <rule_context>epn:VESPA_EPNcore_Parameters/epn:spatial_axes</rule_context>
    <DD_Rule_Statement>
        <rule_type>Assert</rule_type>
        <rule_test>epn:spatial_frame_type = 'none' or
            (epn:spatial_frame_type = 'celestial' and epn:celestial_frame_axes) or
            (epn:spatial_frame_type = 'body' and epn:body_frame_axes) or
            (epn:spatial_frame_type = 'cartesian' and epn:cartesian_frame_axes) or
            (epn:spatial_frame_type = 'spherical' and epn:spherical_frame_axes) or
            (epn:spatial_frame_type = 'sylindrical' and epn:cylindrical_frame_axes)</rule_test>
        <rule_message>The *_frame_axes element corresponding to the spatial_frame_type must be present</rule_message>
    </DD_Rule_Statement>
</DD_Rule>
```

EPNcore LDD option 4
Detailed coords schema

Example_label.xml – section of file, with spatial keywords

```
<spatial_axes>
    <spatial_frame_type>cartesian</spatial_frame_type>
    <spatial_coordinate_description>Jupiter_III</spatial_coordinate_description>
    <spatial_origin>Jupiter</spatial_origin>
    <cartesian_frame_axes>
        <x_min unit="m">10</x_min>
        <x_max unit="m">100</x_max>
        <y_min unit="m">-5</y_min>
        <y_max unit="m">5</y_max>
        <z_min unit="m">100</z_min>
        <z_max unit="m">10000</z_max>
    </cartesian_frame_axes>
    <s_region>
        <s_region_type>polygon</s_region_type>
        <s_region_polygon_point>
            <s_region_lon unit="deg">10</s_region_lon>
            <s_region_lat unit="deg">-5</s_region_lat>
        </s_region_polygon_point>
        <s_region_polygon_point>
            <s_region_lon unit="deg">10</s_region_lon>
            <s_region_lat unit="deg">5</s_region_lat>
        </s_region_polygon_point>
```