PDS4 Local Data Dictionaries

Europa Clipper PDS4 Training Session

June 14, 2018

JPL

Edward Guinness

PDS Geosciences Node
Washington University

Introduction

- Introduce PDS4 dictionary concepts and terminology
- PDS4 Information Model
  - Common (core) dictionary
- Local data dictionaries (LDD)
  - Discipline and Mission dictionaries
- Examples of discipline dictionaries
- Examples of mission dictionary (MESSENGER)
- Process of creating a local data dictionary

PDS4 Dictionary Terminology

- Information Model (IM) – conceptual model for describes metadata objects.
- Data dictionary (or dictionary) – defines metadata attributes and their relationships within a domain.
- PDS4 common (core) dictionary – metadata derived from the IM and covers.
- Local data dictionary (LDD) – a data dictionary that deals within a limited domain.
  - Discipline dictionary addresses a particular type of data or discipline (images, spectra, geometry, cartography, etc.).
  - Mission dictionary provides a local vocabulary for a specific mission or investigation.

PDS4 Information Model

- Defines explicit relationships between major entities of PDS.
- Establishes governance for PDS4 metadata for the core dictionary and is extended to discipline and mission level dictionaries.
- Provides consistency in PDS labels across many instruments and observation types.
- Provides a single authoritative source for PDS data standards.
- Based on international standard for data dictionaries.
- The PDS4 IM is independent of implementation.
  - Currently the IM is implemented in XML for labels.

XML Use in PDS4

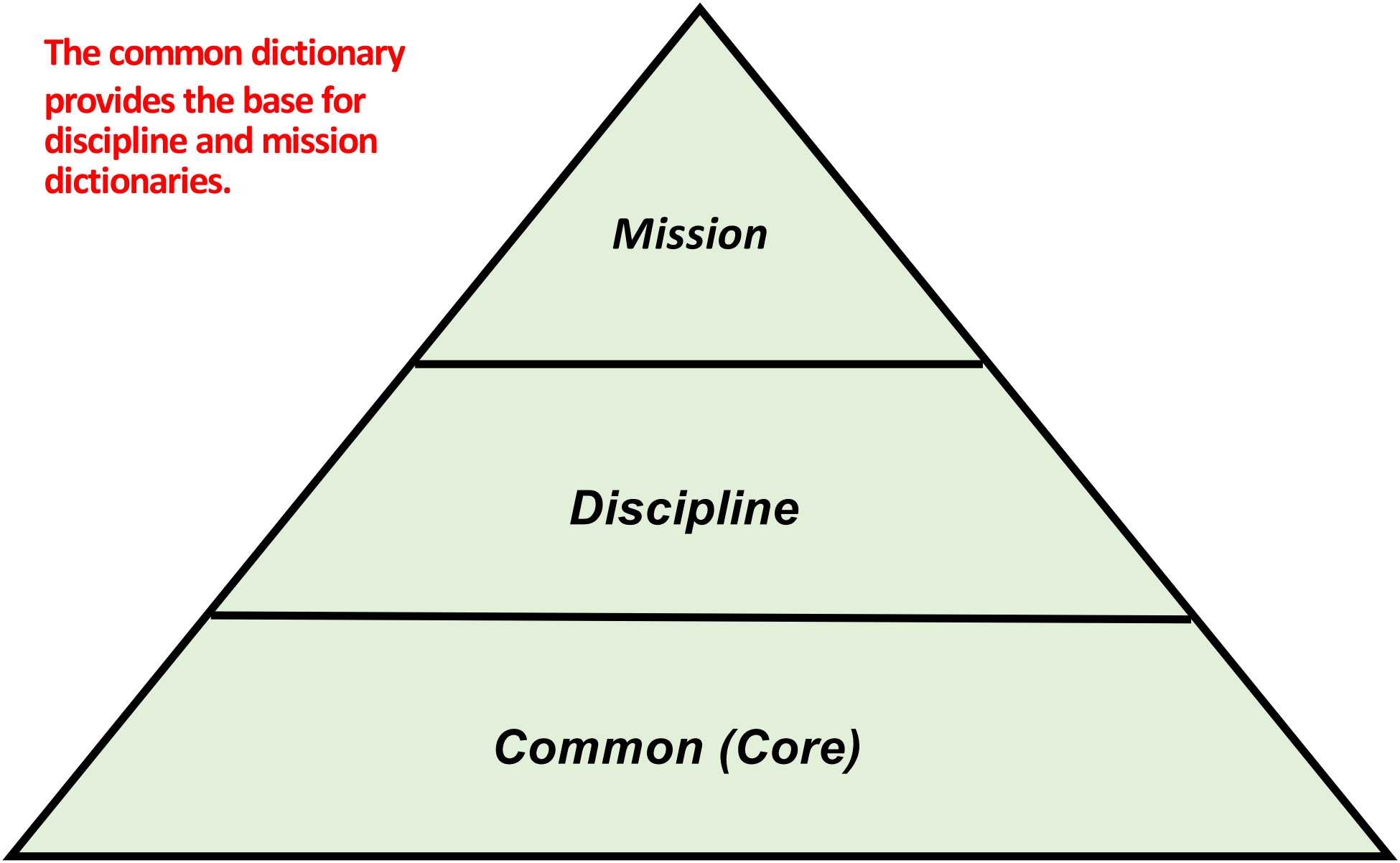
- All PDS4 dictionaries are represented as XML schema and schematron.
  - Each dictionary has a schema file and schematron file.
- XML schema is used to:
  - Define metadata attributes and groups of attributes (classes);
  - Specify their order in a label;
  - Constrain attribute content.
- Schematron provides additional rules for constraining label content.
  - Standard value lists.
  - Enforce context-dependent constraints (“If this, then that”; “Either this or that, but not both”).

PDS4 Dictionary Development

- PDS4 common dictionary is relatively stable.
  - Changes must be approved by a Change Control Board.
  - Updates, if needed, are released every 6 months and are version controlled.
  - All versions are archived and available at PDS site.
- Local dictionaries (Discipline and Mission) are locally controlled and more dynamic.
  - However, they are still version controlled and all versions are available at the PDS site.
  - Local dictionaries can reference elements from the common dictionary – but not the other direction.

PDS4 Dictionary Relationships

The common dictionary provides the base for discipline and mission dictionaries.



Examples of Discipline Dictionaries

- Display Dictionary
  - Used to define how array (image) data should be displayed on a display device.
- Geometry Dictionary
  - Defines the observational geometry relevant to data objects in the product.
- Cartography Dictionary
  - Defines the attributes to describe the map projections in cartographic products.
- Rings Dictionary
  - Defines supplemental metadata that support ring system observations.

Display Dictionary

- Associates horizontal and vertical display directions with axes of an array.
- Defines direction that lines and samples are to be drawn (top-to-bottom, left-to-right).

```
<disp:Display_Settings>
 <disp:Local_Internal_Reference>
  <local_identifier_reference>Spectral_Qube_Object</local_identifier_reference>
  <local_reference_type>display_settings_to_array</local_reference_type>
 </disp:Local_Internal_Reference>
 <disp:Display_Direction>
  <disp:horizontal_display_axis>Sample</disp:horizontal_display_axis>
  <disp:horizontal_display_direction>Left to Right</disp:horizontal_display_direction>
  <disp:vertical_display_axis>Line</disp:vertical_display_axis>
  <disp:vertical_display_direction>Top to Bottom</disp:vertical_display_direction>
 </disp:Display_Direction>
</disp:Display_Settings>
```

Geometry Dictionary

- Provides metadata for the geometric conditions with which observational data were collected.
  - Lighting and viewing angles.
  - Distances between relevant objects, such as Sun to target, spacecraft to target.
  - Position and velocity vectors of the observer relative to the target being observed.
  - Location and orientation of an observation projected onto the surface of the target.
  - Links to SPICE kernels used to derive geometry parameters.
- Has classes for orbital/flyby and landed geometry.

Cartography Dictionary

- Used to describe parameters associated with map projected data products.
- Based on a federal geographic data standard for describing map projections.
- Has classes for:
  - Defining a bounding box for the limits of coverage of a map;
  - Supports many map projections (e.g., cylindrical, equirectangular, Mercator, polar stereographic, etc.).

Mission Dictionary

- Currently, PDS4 mission dictionaries are available for Maven and InSight.
- Several additional mission dictionaries to support PDS3 to PDS4 migrations are in development.
- A mission dictionary might includes parameters for:
  - Mission phases;
  - Spacecraft clocks;
  - Other mission specific times (orbit number, surface times);
  - Engineering parameters for instruments.

MESSENGER mission dictionary

```
<Mission_Area>
  <mess:MESSENGER>
   <mess:mission_phase_name>Mercury Orbit Year 5</mess:mission_phase_name>
   <mess:spacecraft_clock_start_count>2/070170410</mess:spacecraft_clock_start_count>
   <mess:spacecraft_clock_stop_count>2/070231471</mess:spacecraft_clock_stop_count>
   <mess:standard_data_product_id>ns_calibration_diagnostic</mess:standard_data_product_id>
   <mess:software_name>pipe-ns2edr</mess:software_name>
   <mess:software_version_id>1.0</mess:software_version_id>
  </mess:MESSENGER>
 </Mission_Area>
```

How to build a discipline or mission dictionary

- Always good to consult with a PDS discipline node when building a LDD.
- Mission dictionary coordinated by PDS lead node for the mission.
  - For Europa Clipper the Imaging Node is the lead node.
- Could have one mission dictionary that also contains classes for instruments.
- Could have one mission dictionary and separate instrument dictionaries.
  - Instrument dictionaries would be coordinated with the PDS node that will archive that instrument's data.

Ingest_LDD

- An LDD starts with the creation and population of an XML file known as an ingest_LDD file.
- Ingest_LDD is a class with the PDS4 core dictionary.
- It has methods for defines individual keywords (attributes) and groups of keywords (classes).
- It is also where constraints (rules) for schematron are defined.

Examples from Ingest_LDD - Attribute

```
<DD_Attribute>
        <name>mission_phase_name</name>
        <version_id>1.0</version_id>
        <local_identifier>mess.mission_phase_name</local_identifier>
        <nillable_flag>false</nillable_flag>
        <submitter_name>PDS3 Migration</submitter_name>
        <definition>The mission_phase_name attribute provides the mission-defined
            name of a mission phase.</definition>
        <DD_Value_Domain>
            <enumeration_flag>true</enumeration_flag>
            <value_data_type>ASCII_Short_String_Collapsed</value_data_type>
            <maximum_characters>255</maximum_characters>
            <DD_Permissible_Value>
              <value>Launch</value>
              <value_meaning>2004-08-03 (2004-216)to 2004-09-12 (2004-256)</value_meaning>
            </DD_Permissible_Value>
            <DD_Permissible_Value>
              <value>Earth Cruise</value>
              <value_meaning>2004-09-13 (2004-257) to 2005-07-18 (2005-199)</value_meaning>
            </DD_Permissible_Value>
            <DD_Permissible_Value>
              <value>Earth Flyby</value>
              <value_meaning>2005-07-19 (2005-200) to 2005-08-16 (2005-228)</value_meaning>
            </DD_Permissible_Value>
.
.
      </DD_Value_Domain<
</DD_Attribute
```

Examples from Ingest_LDD - Class

```
<DD_Class>
        <name>MESSENGER</name>
        <version_id>1.0</version_id>
        <local_identifier>mess.mdis.MESSENGER</local_identifier>
        <submitter_name>PDS3 Migration</submitter_name>
        <definition>The MESSENGER class is the container for MESSENGER mission-specific metadata elements</definition>
        <element_flag>true</element_flag>
        <DD_Association>
            <identifier_reference>mess.mission_phase_name</identifier_reference>
            <reference_type>attribute_of</reference_type>
            <minimum_occurrences>0</minimum_occurrences>
            <maximum_occurrences>1</maximum_occurrences>
        </DD_Association>
        <DD_Association>
            <identifier_reference>mess.spacecraft_clock_start_count</identifier_reference>
            <reference_type>attribute_of</reference_type>
            <minimum_occurrences>0</minimum_occurrences>
            <maximum_occurrences>1</maximum_occurrences>
        </DD_Association>
        <DD_Association>
            <identifier_reference>mess.spacecraft_clock_stop_count</identifier_reference>
            <reference_type>attribute_of</reference_type>
            <minimum_occurrences>0</minimum_occurrences>
            <maximum_occurrences>1</maximum_occurrences>
        </DD_Association>
        .
        .
        .
</DD_Class>
```

Examples from Ingest_LDD - Rules

```
<DD_Rule_Statement>
    <rule_type>Assert</rule_type>
    <rule_test>if ( ((insight:start_sol_number) and (not(insight:stop_sol_number))) or
      ((insight:stop_sol_number) and (not(insight:start_sol_number))) ) then false() else true() </rule_test>
    <rule_message>insight:error:sol_number_1: If either of the pair insight:start_sol_number and
      insight:stop_sol_number is present, then both must be present.</rule_message>
    <rule_description>If either of the pair start_sol_number and stop_sol_number is present, then
      both must be present.</rule_description>
  </DD_Rule_Statement>
```

This rule ensures that if either start_sol_number or stop_sol_number is in the label, then both must be present. In other words, one cannot be used without the other.

DD_Rule_Statement is converted into schematron rule by the LDD creation tool.

Building a LDD with LDDtool

- PDS provides a command line tool that interprets an Ingest_LDD file and outputs an XML schema and schematron file.
- LDDtool does some validation to make sure the LDD is consistent with the core dictionary.
- LDDtool can output the LDD in other formats.
  - For example, the dictionary can be output as a JSON file.
  - Useful for integrating the dictionary into an archive production pipeline.

Concluding Remarks

COMMUNICATION!

Start working with the PDS Discipline Node that will be curating your archive early in the process and communicate with them regularly!

Work closely with that node and the lead node in developing the required local dictionaries.

Dictionary development is likely to be an iterative process.