Model-Driven Development For PDS4 Software And Services

Steven Hughes, Daniel Crichton, Stirling Algermissen, Michael Cayanan, Ronald Joyner, Sean Hardman, and Jordan Padams

NASA Jet Propulsion Laboratory (JPL), California Institute of Technology

Planetary Science Informatics and Data Analytics (PSIDA)
Washington University, St. Louis, MO - April 24-26, 2018

Data Architectures, Management, And Data Technologies
Tuesday, April 24, 2018 - 12:05 pm
Topics

- Overview of Model-Driven Development
- Benefits to the Information System
- Current Applications
- Conclusion
Information Model

• “An information model is a representation of concepts, relationships, constraints, rules, and operations to specify data semantics for a chosen domain of discourse.”

• It provides a sharable, stable, and organized structure of **information requirements** or knowledge for the domain context.

---

Information Model

Planetary Science Domain Expertise → PDS4 Information Model → Extract Filter Translate → XML Schema And Schematron

PDS4 Information Model

PDS4 System Requirements

JSON XMI/UML RDF/XML OWL/XML SKOS/XML

Data Dictionary Document

Information Model Specification

Registry Configuration
Background

• Developed using lessons learned from over 25 years of science data archiving
  – Acquired a good understanding of data representations in different planetary disciplines

• Used best practices for information model development and foundational principles adopted from:
  – Open Archival Information System (OAIS) Reference Model - ISO 14721 - Foundational Principles
  – CCSDS 312.0-G-1 - Reference Architecture for Space Information (RASIM)
  – Management - W3C XML (Extensible Markup Language) - Rules for encoding documents electronically

• Few systems have developed a rigorous model for describing data management, discovery, and analytics
Community’s View

Information Modeler’s View

Repository View

Product

Tagged Data Object
(Information Object)

<local_identifier>MPFL_M_IMP_IMAGE</local_identifier>
<offset unit="byte">0</offset>
<axes>2</axes>
<axis_index_order>Last_Index_Fastest</axis_index_order>
<encoding_type>Binary</encoding_type>
<Element_Array>
<data_type>SignedMSB4</data_type>
<unit>pixel</unit>
</Element_Array>
<Axis_Array>
<axis_name>Line</axis_name>
<elements>248</elements>
<sequence_number>1</sequence_number>
</Axis_Array>
<Axis_Array>
<axis_name>Sample</axis_name>
<elements>256</elements>
<sequence_number>2</sequence_number>
</Axis_Array>
</Array_2D_Image>
Roles of the IM

- Captures domain expertise:
  - *science interpretation and use of the data*
  - *context within which the data was captured, processed, and archived*

- Defines:
  - *data structure (format)*
  - *relationships between the data*

- Single authoritative source for the data standards
  - *Promotes a self-describing archive*

- Drives the PDS4 infrastructure by providing:
  - *A sharable, stable, and organized structure of information requirements.*
  - *Formal definitions that are suitable for configuring and generating code.*

- Remains independent of the implementation
ISO¹ Standard Governance Entities

- Registration Authority
- Steward
- Namespace

¹ ISO 14721:2003 - Open Archival Information System (OAIS) Reference Model
Benefits - Interoperability

- Enables the development of software and services to support interoperability
  
  - *The Common dictionary provide terminology (syntactic and semantic) that enables interoperability across the entire community*
  
  - *The Discipline dictionaries enable interoperability within science and engineering disciplines*
    - Cartography and Geometry
    - Rings and Atmospheres
  
  - *The Mission dictionaries provide a local vocabulary for a mission or project.*
Minimizes the Impact of Change

• An independent Information Model disentangles the information model from the implementation technology.
  – *Technology changes at a rate different from the domain*
  – *Software and Services can be designed to respond to the information model*

• Multi-level governance limits the impact of change
  – *The Common dictionary is relatively stable*
  – *The Mission and Project dictionaries are localized and more dynamic.*
Extensions

- Extensions to the model inherit the full capability of the parent model, while retaining the ability to add customized capability.
  - The Common dictionary defines the core entities: products, collections of products, data types, and units of measure
  - The Discipline and Mission reference the common elements as necessary
A Self Describing Archive

<table>
<thead>
<tr>
<th>Identification_Area</th>
<th>Logical_Identifier</th>
<th>Version_Id</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Observation_Area</th>
<th>Time_Coordinates</th>
<th>Primary_Result_Summary</th>
<th>Investigation_Area</th>
<th>Observing_System</th>
<th>Target_Identification</th>
</tr>
</thead>
</table>

| Discipline_Area      | Mission Area       |

| Reference_List       | Internal_Reference | External_Reference     |

| File_Area_Observational | File | Header | Array_2D_Image |

June 8, 2015
<Product_Observational>
  <Identification_Area>
    <logical_identifier>urn:nasa:pds:example.dph.sampleproducts:exampleproducts:array2d_image ...
    <version_id>1.0</version_id>
    <title>MARS PATHFINDER LANDER Experiment</title>
  </Identification_Area>
  <Array_2D_Image>
    <local_identifier>MPFL-M-IMP_IMG_GRAYSCALE</local_identifier>
    <offset unit="byte">0</offset>
    <axes>2</axes>
    <axis_index_order>Last Index Fastest</axis_index_order>
    <Element_Array>
      <data_type>UnsignedMSB2</data_type>
      <unit>data number</unit>
      <scaling_factor>1</scaling_factor>
      <value_offset>0</value_offset>
    </Element_Array>
    <Axis_Array>
      <axis_name>Line</axis_name>
      <elements>248</elements>
      <sequence_number>1</sequence_number>
    </Axis_Array>
    <Axis_Array>
      <axis_name>Sample</axis_name>
      <elements>256</elements>
      <sequence_number>2</sequence_number>
    </Axis_Array>
  </Array_2D_Image>
</Product_Observational>
The Array class defines a homogeneous N-dimensional array of scalars. The Array 2D Image class is an extension of the Array 2D class and defines a two dimensional image.

The attribute `pds:axis_index_order` must be equal to the value 'Last Index Fastest'.

The attribute `pds:axis_index_order` must be equal to the value 'Last Index Fastest'.

The attribute `pds:axis_index_order` must be equal to the value 'Last Index Fastest'.

The attribute `pds:axis_index_order` must be equal to the value 'Last Index Fastest'.

The attribute `pds:axis_index_order` must be equal to the value 'Last Index Fastest'.

The attribute `pds:axis_index_order` must be equal to the value 'Last Index Fastest'.

The attribute `pds:axis_index_order` must be equal to the value 'Last Index Fastest'.
Validation

Lines of XML Schema and Schematron

Common

Discipline

- Common
- Display
- Rings
- Cartography
- Geometry
- Imaging
Terminological_Entry

<identifier>pds3:ARM_ARTICULATION_STATE.ARTICULATION_DEVICE_ANGLE_NAME</identifier>
<namespace_id>pds3</namespace_id>
<steward_id>pds3</steward_id>
	<title>ARM_ARTICULATION_STATE.ARTICULATION_DEVICE_ANGLE_NAME</title>

<referenced_identifier>
 insight:Instrument_Parameters  insight:index_value_angle
</referenced_identifier>

<skos_relation_name>closeMatch</skos_relation_name>

<instance_id>npds:Observation_Area/pds:Discipline_Area/geom:Geometry/geom:Geometry_Lander/geom:Articulation_Device_Parameters[1]/geom:Device_Angle/geom:Device_Angle_Index/geom:index_value_angle</instance_id>
Analytics

Identify data products within the Planetary Data System (PDS) Archive that are scientifically useful for the Exoplanet project.

Number Of Instruments per Energy/Wavelength Facet

- visible: 152
- microwaves: 54
- infrared: 33
- near-infrared: 29
- magnetic-field: 19
- ultraviolet: 12
- x-ray: 9
- thermal-infrared: 7
- mass: 7
- gamma-ray: 6
- radio: 3
- electrons: 3
- ions: 3
- keV: 2
- neutrals-plasma: 2
- swir: 2
- MeV: 1
- energetic: 1
- cosmic ray: 1
- charged particles: 1

0 20 40 60 80 100 120 140 160
1 1 1 1 1 1 2 2 3 3 3 6 7 7 9 12 19 22 29 33 54
Conclusion

- The PDS4 Information Model is the core of the PDS4 Information System.
  - Provides the Information Requirements for the system.
  - Used to help configure common services and software
  - Provides the basis for a self-describing archive

- The semantic and syntactic information in the model is increasingly used to support data analytics

- Being used as a prototype in the development of an implementable architecture for Trusted Digital Repositories.
  - Open Archival Information System (OAIS) Reference Model – ISO-14721
Thank You

Questions and Answers

PDS homepage: https://pds.nasa.gov/

Acknowledgements – Special thanks to Cristina De Cesare for her support on Term Mapping.

Acknowledgements - This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.
Backup
Core Components in Context

Internet Users

Internet Service Endpoints
Search, Retrieve, Transport, Transform, ...

The PDS

Product Centric

External Resources

Curated Data

Discipline

Product

Context

Instrument_Host

Instrument

Observing_System

Node

Target

Investigation

Bundle

Collection

Product_Observational

has_member

Document

Array

Array_2D_Image

has

has_member

part_of

defines

Digital_Object

is-a

is-a

is-a

is-a

is-a

is-a

is-a

composed_of

composed_of

curated_by

is
<table>
<thead>
<tr>
<th>Registration Authority</th>
<th>Steward Id</th>
<th>Namespace Id*</th>
<th>XML Schema Namespace</th>
<th>Logical Identifier Prefix</th>
<th>Governance Level</th>
<th>Steward</th>
<th>Oversight</th>
</tr>
</thead>
<tbody>
<tr>
<td>0001_NASA_PDS_1</td>
<td>pds</td>
<td>pds</td>
<td><a href="http://pds.nasa.gov/pds4/pds/v1">http://pds.nasa.gov/pds4/pds/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Common</td>
<td>PDS EN Node</td>
<td>CCB</td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>atm</td>
<td>atm</td>
<td><a href="http://pds.nasa.gov/pds4/atm/v1">http://pds.nasa.gov/pds4/atm/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Discipline</td>
<td>PDS ATM Node</td>
<td></td>
</tr>
<tr>
<td>0001_JAXA_DARTS_1</td>
<td>darts</td>
<td>darts</td>
<td><a href="http://pds.nasa.gov/pds4/darts/v1">http://pds.nasa.gov/pds4/darts/v1</a></td>
<td>urn:jaxa:darts:</td>
<td>Discipline</td>
<td>DARTS (JAXA)</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>en</td>
<td>dph</td>
<td><a href="http://pds.nasa.gov/pds4/dph/v1">http://pds.nasa.gov/pds4/dph/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Discipline</td>
<td>PDS EN Node</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>geo</td>
<td>geo</td>
<td><a href="http://pds.nasa.gov/pds4/geo/v1">http://pds.nasa.gov/pds4/geo/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Discipline</td>
<td>PDS GEO Node</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>img</td>
<td>cart</td>
<td><a href="http://pds.nasa.gov/pds4/cart/v1">http://pds.nasa.gov/pds4/cart/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Discipline</td>
<td>PDS IMG Node</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>img</td>
<td>disp</td>
<td><a href="http://pds.nasa.gov/pds4/disp/v1">http://pds.nasa.gov/pds4/disp/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Discipline</td>
<td>PDS IMG Node</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>naif</td>
<td>naif</td>
<td><a href="http://pds.nasa.gov/pds4/naif/v1">http://pds.nasa.gov/pds4/naif/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Discipline</td>
<td>PDS NAIF Node</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>ops</td>
<td>pds</td>
<td><a href="http://pds.nasa.gov/pds4/pds/v1">http://pds.nasa.gov/pds4/pds/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Discipline</td>
<td>PDS EN Node</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>ppi</td>
<td>alt</td>
<td><a href="http://pds.nasa.gov/pds4/alt/v1">http://pds.nasa.gov/pds4/alt/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Discipline</td>
<td>PDS PPI Node</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>ppi</td>
<td>particle</td>
<td><a href="http://pds.nasa.gov/pds4/particle/v1">http://pds.nasa.gov/pds4/particle/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Discipline</td>
<td>PDS PPI Node</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>ppi</td>
<td>ppi</td>
<td><a href="http://pds.nasa.gov/pds4/ppi/v1">http://pds.nasa.gov/pds4/ppi/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Discipline</td>
<td>PDS PPI Node</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>ppi</td>
<td>wave</td>
<td><a href="http://pds.nasa.gov/pds4/wave/v1">http://pds.nasa.gov/pds4/wave/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Discipline</td>
<td>PDS PPI Node</td>
<td></td>
</tr>
<tr>
<td>0001_ESA_PSA_1</td>
<td>psa</td>
<td>psa</td>
<td><a href="http://psa.esa.int/psa/v1">http://psa.esa.int/psa/v1</a></td>
<td>urn:psa:esa:</td>
<td>Discipline</td>
<td>ESA PSA</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>rings</td>
<td>rings</td>
<td><a href="http://pds.nasa.gov/pds4/rings/v1">http://pds.nasa.gov/pds4/rings/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Discipline</td>
<td>PDS Rings Node</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>rs</td>
<td>rs</td>
<td><a href="http://pds.nasa.gov/pds4/rs/v1">http://pds.nasa.gov/pds4/rs/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Discipline</td>
<td>PDS RS Node</td>
<td></td>
</tr>
<tr>
<td>0001_ROS_RSSA_1</td>
<td>rssa</td>
<td>rssa</td>
<td><a href="http://pds.nasa.gov/pds4/rssa/v1">http://pds.nasa.gov/pds4/rssa/v1</a></td>
<td>urn:ros:rssa:</td>
<td>Discipline</td>
<td>RSSA (IKI)</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>sbn</td>
<td>sbn</td>
<td><a href="http://pds.nasa.gov/pds4/sbn/v1">http://pds.nasa.gov/pds4/sbn/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Discipline</td>
<td>PDS SBN</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>sbn</td>
<td>sp</td>
<td><a href="http://pds.nasa.gov/pds4/sp/v1">http://pds.nasa.gov/pds4/sp/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Discipline</td>
<td>PDS SBN</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>atm</td>
<td>ladee</td>
<td><a href="http://pds.nasa.gov/pds4/ladee/v1">http://pds.nasa.gov/pds4/ladee/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Mission</td>
<td>PDS ATM Node</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>ppi</td>
<td>mvn</td>
<td><a href="http://pds.nasa.gov/pds4/mission/mvn/v1">http://pds.nasa.gov/pds4/mission/mvn/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Mission</td>
<td>PDS PPI Node</td>
<td></td>
</tr>
<tr>
<td>0001_NASA_PDS_1</td>
<td>ppi</td>
<td>mvn</td>
<td><a href="http://pds.nasa.gov/pds4/mvn/v1">http://pds.nasa.gov/pds4/mvn/v1</a></td>
<td>urn:nasa:pds:</td>
<td>Mission</td>
<td>PDS PPI Node</td>
<td></td>
</tr>
</tbody>
</table>
The PDS4 Information Model

Knowledge Acquisition

Information Requirements

Domain Knowledge

Open Archive Information System Reference Model (ISO 14721)

Data Dictionary Reference Model (ISO/IEC 11179)

Federated Registry Reference Model (ebXML)

Information Model

Information-Base

Ontology Modeling Tool

Protégé

Transform Modules

Extract
Filter
Translate

Documentation, Specification, Requirements, and Guidance

XML Schema & Schematron

Registry Configuration Parameters

XML Documents (Label Templates)

Query Models

Information Model Specification

XMI/UML

RDF/OWL

JSON

SKOS

Data Dictionary Document
UML Class Diagram