



PDS4 Operational Readiness Review  
PDS4 Information Model and Standards Support  
for LADEE/MAVEN

S. Hughes  
Engineering Node

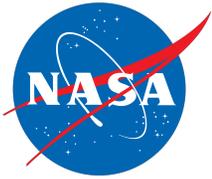




# Topics



- Introduction
- Key Requirements
- Build 3b Deliverables
- Change Requests for Version 1.0
- Conclusion



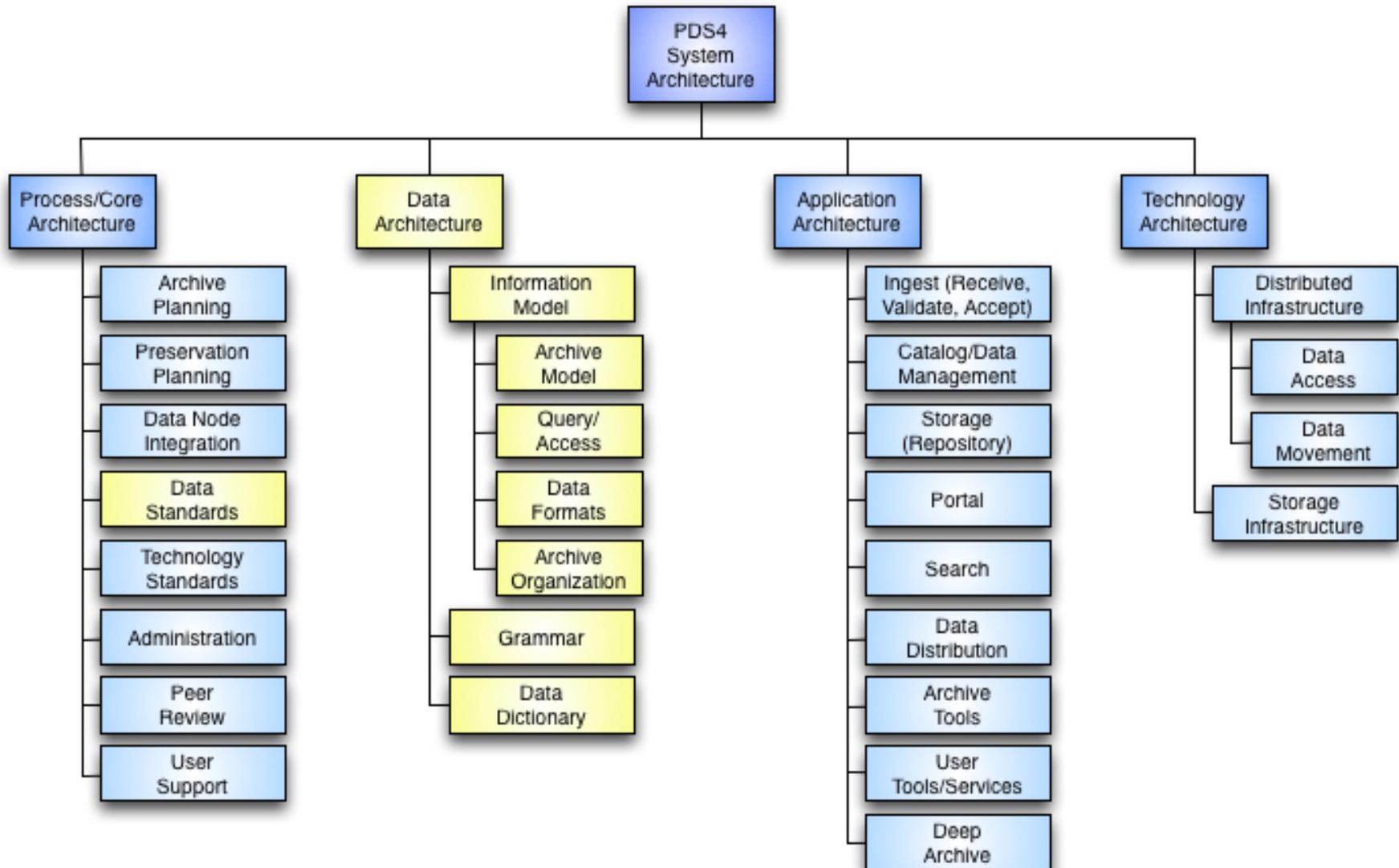
# Introduction

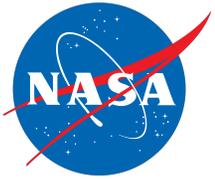
- PDS has developed a new set of standards for PDS4 in support of this PDS Level 2 requirement:
  - 1.4 Archiving Standards: PDS will have archiving standards for planetary science data
- PDS4 has defined a common planetary information model as the “heart” of the PDS4 standards.
- In geographical distributed data systems that involves highly diverse science disciplines and heterogeneous data types, a shared information model is key to enabling system interoperability and data correlation.





# PDS4 Architecture





# The Purpose of the Information Model



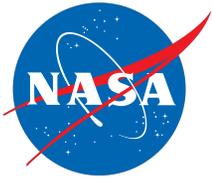
- Defines the data structure (format)
- Defines the science interpretation of the data
- Defines the context within which the data was captured, processed, and archived
- Defines the relationships between the data



# The Design Principles of the Information Model



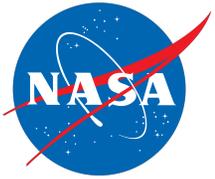
- The information model should remain independent of its implementation.
  - Disentangles the model from the implementation
  - Information model evolves independent of information technology
- A changing domain suggests that the information model should drive both the development and management of the information system.
- The modeling language should be semantically richer than the other languages in the framework.



# Knowledge Acquisition for the Information Model



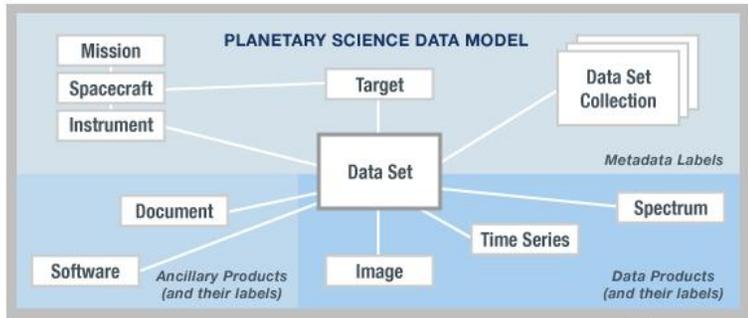
- Domain expertise is captured in an ontology.
  - A working group was formed with at least one domain expert from each of the science disciplines.
  - Each thing-of-interest in the domain was defined and then related to other things-of-interest.
  - The resulting model represents the consensus of domain experts across the PDS science and engineering disciplines.
- The model is subsequently used as the single authoritative source for the PDS4 Data Standards.



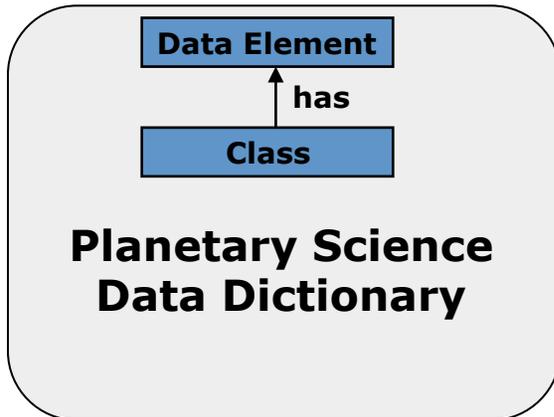
# The Primary Role of the Information Model



## Information Model



Expressed As



Used to Create

Validates



Extracted/Specialized

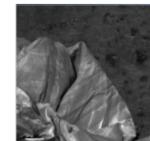
## Product

### Tagged Data Object

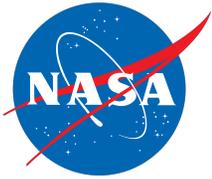
(Information Object)

```
<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>
```

Describes



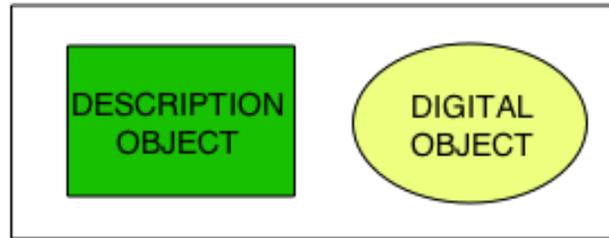
Data Object



# Information Object Model <sup>1</sup>

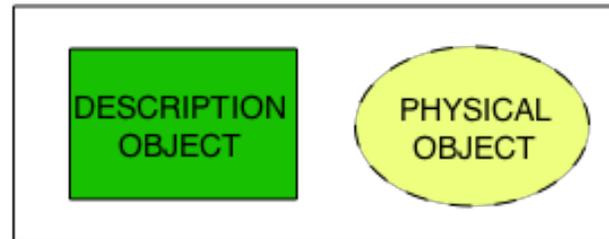


TAGGED DIGITAL OBJECT



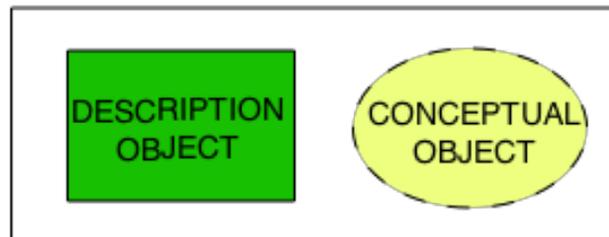
- **digital object:** An object which is real data — for example, a binary image of a redwood tree.

TAGGED NON-DIGITAL OBJECT



- **physical object:** An object which is physical or tangible – for example the planet Saturn and the Venus Express magnetometer.

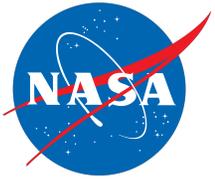
TAGGED NON-DIGITAL OBJECT



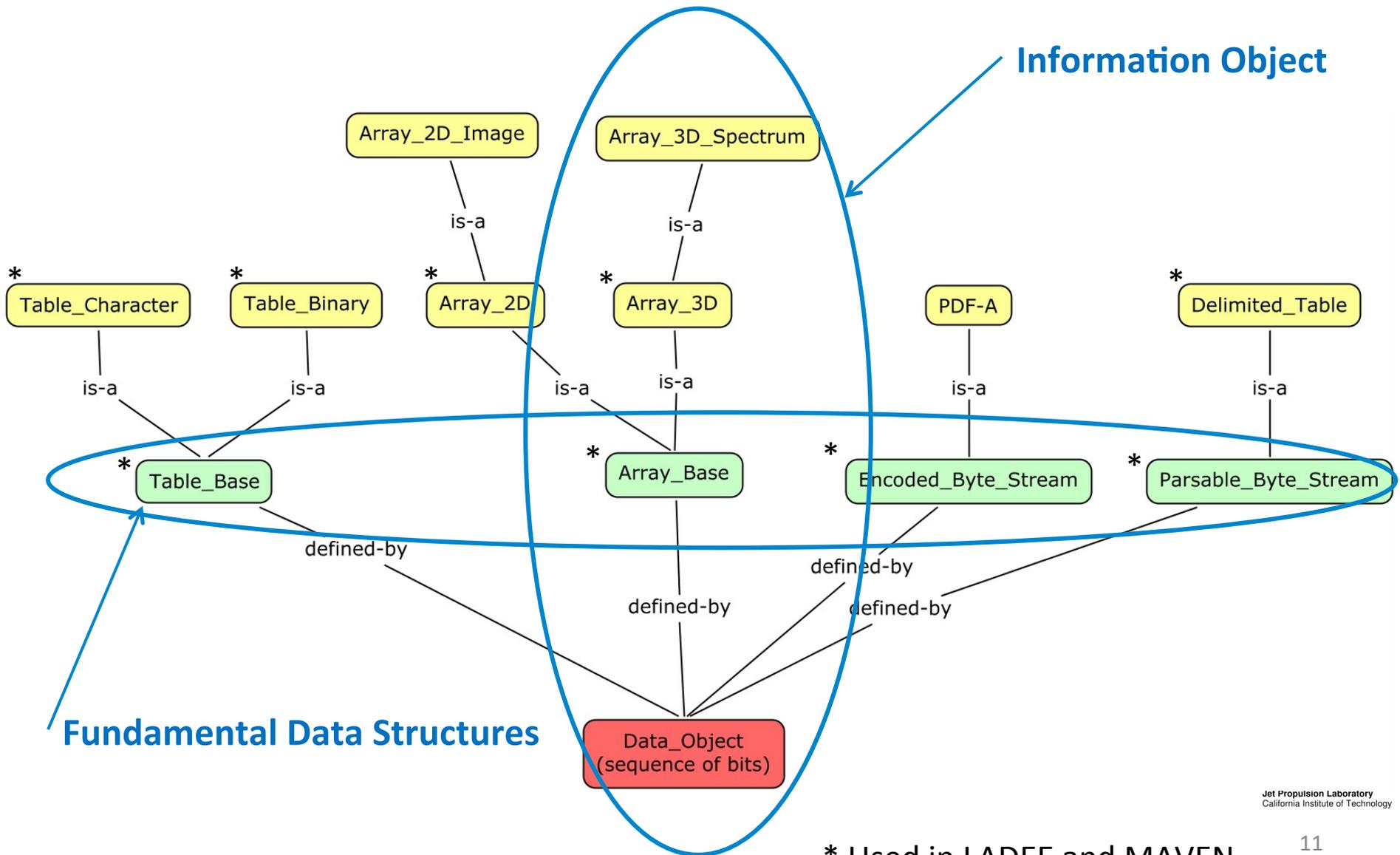
- **conceptual object:** An object which is intangible – for example the Cassini mission and NASA's strategic plan for solar system exploration.

<sup>1</sup> Open Archival Information System (OAIS) Reference Model - ISO 14721:2003



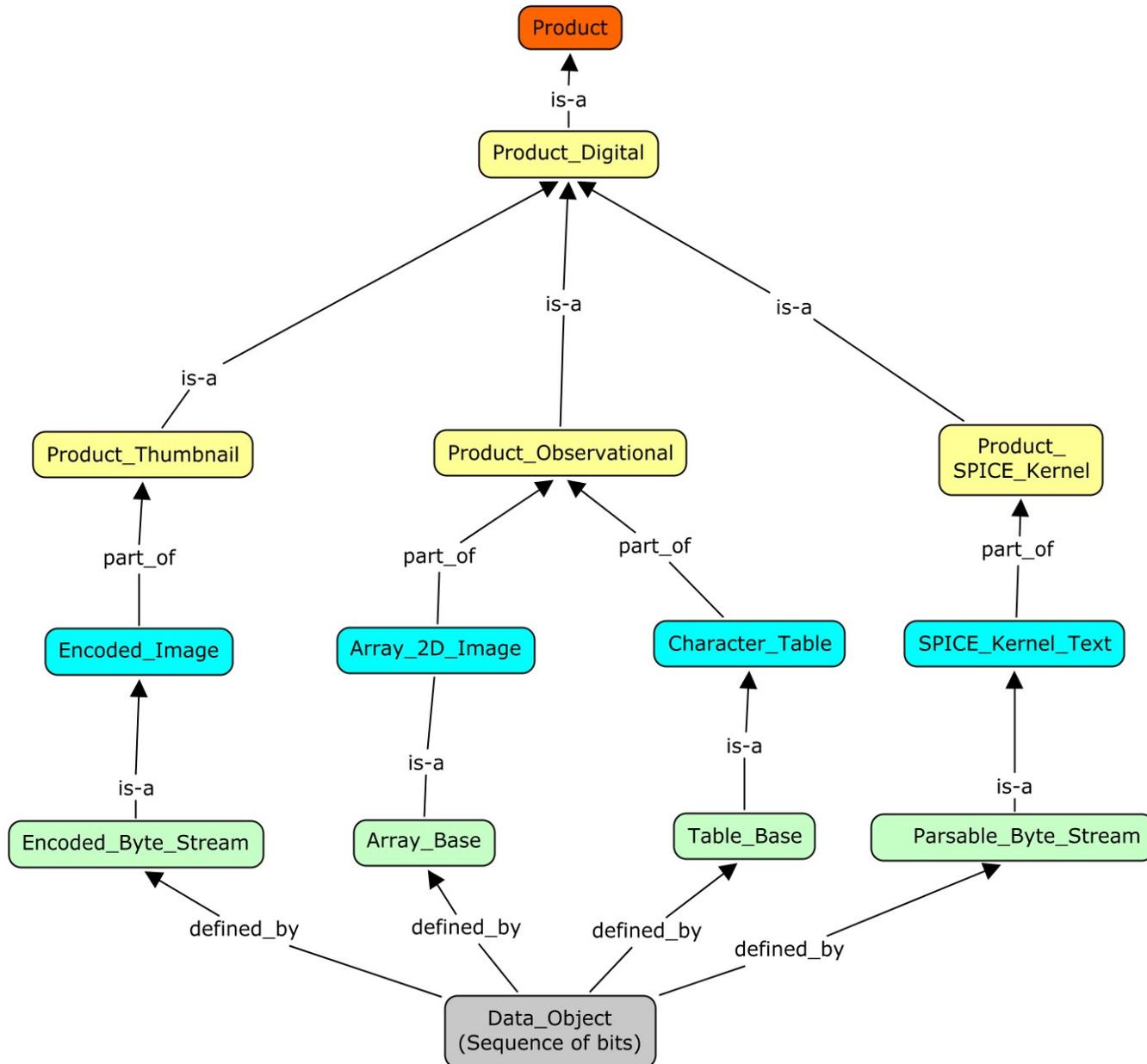


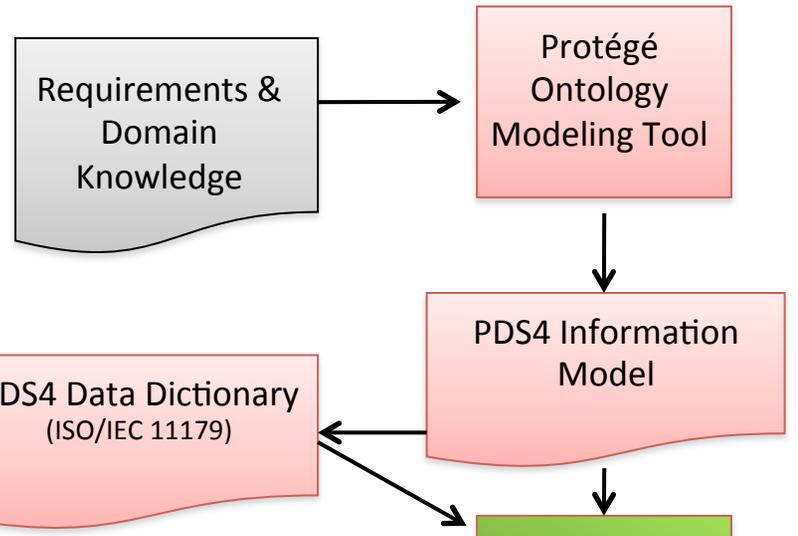
# PDS4 Data Formats



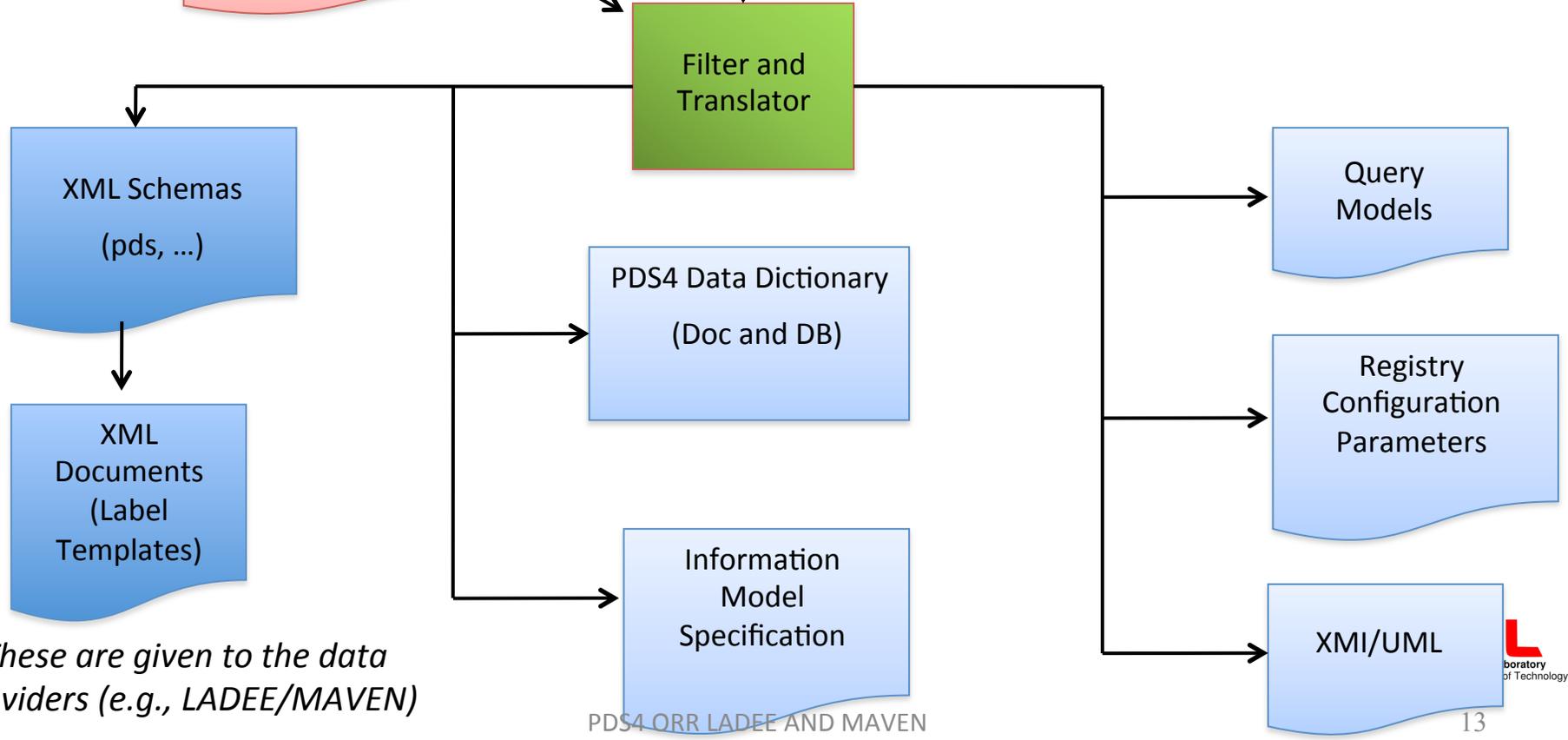


# PDS4 Product Model



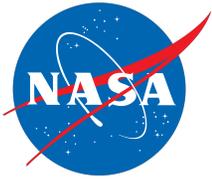


# The Information Model Driven Process & The Generated Artifacts

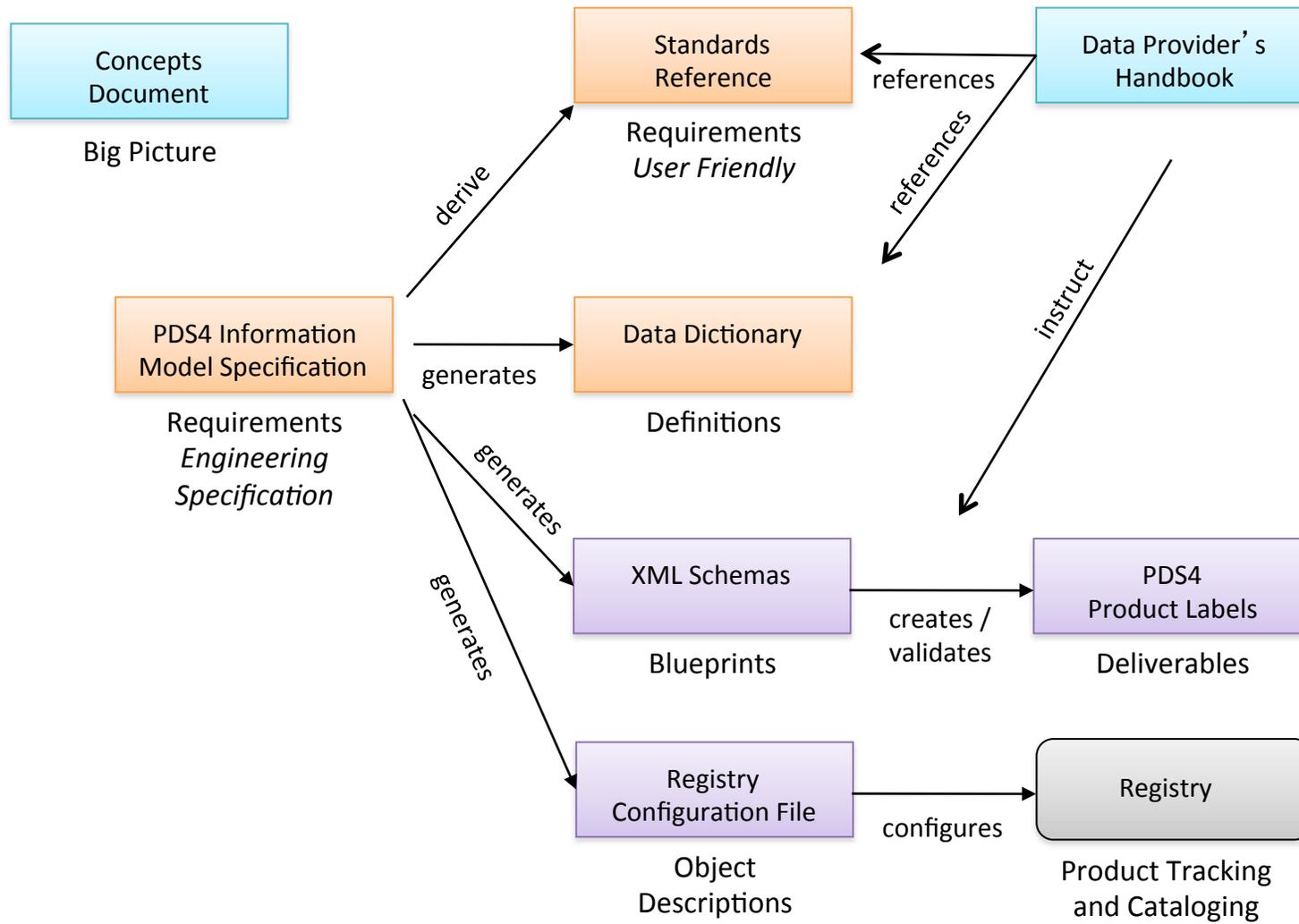


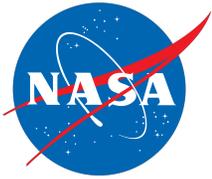
*These are given to the data providers (e.g., LADEE/MAVEN)*





# PDS4 Documents, Artifacts, and their Relationships

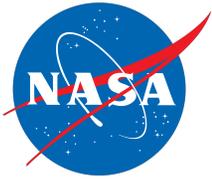




# Product Label Template



Identification_Area	
Logical_Identifier	
Version_Id	
Observation_Area	
Time_Coordinates	<i>Discipline_Area</i>
Primary_Result_Summary	<i>Mission Area</i>
Investigation_Area	
Observing_System	
Target_Identification	
Reference_List	
Internal_Reference	
External_Reference	
File_Area_Observational	
File	
Header	
Array_2d_Image	



# Industry Standards\* Referenced and Controlling



- ISO 14721:2003 - Open Archival Information System (OAIS) Reference Model - Provides a standard for information objects.
- ISO/IEC 11179:3 Registry Metamodel and Basic Attributes specification - Adopted for the data dictionary schema.
- ISO/IEC 11404:2007(E) - Provides the specification for language-independent data types.
- Reference Architecture for Space Information Management (RASIM) - CCSDS 312-0.G-1 – Provides the overarching architectural principles.
- W3C XML (Extensible Markup Language) - Rules for encoding documents electronically.
- W3C XML schema - Type description language for XML documents.
- Electronic Business XML (ebXML) federated registry/repository information model – Provides a standard to support federated registry/repository functions
- RDF/RDFS/XML - RDF is a standard model for data interchange on the Web



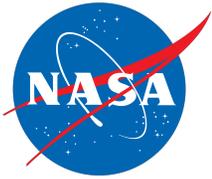
\* Not a complete list



# Data Dictionary Model ISO/IEC 11179



- **Data Element**
  - Name
  - Submitter, **Steward**
  - Definition
  - **Namespace**
  - Source of definition
  - Change log
  - **Version**
  - Concept
  - **Alternate Names**
  - **Definition in multiple natural languages**
  - Classification
  - **Unit of measurement**
  - **Effective Dates**
- **Valid Value**
  - Value
  - Submitter, Steward
  - Definition
  - Cardinality
  - Source of definition
  - Change log
  - Version
  - Concept
  - Character Set
  - Representation
  - Minimum and Maximum Value
  - Minimum and Maximum Length
  - Alternate encodings
  - Effective Dates



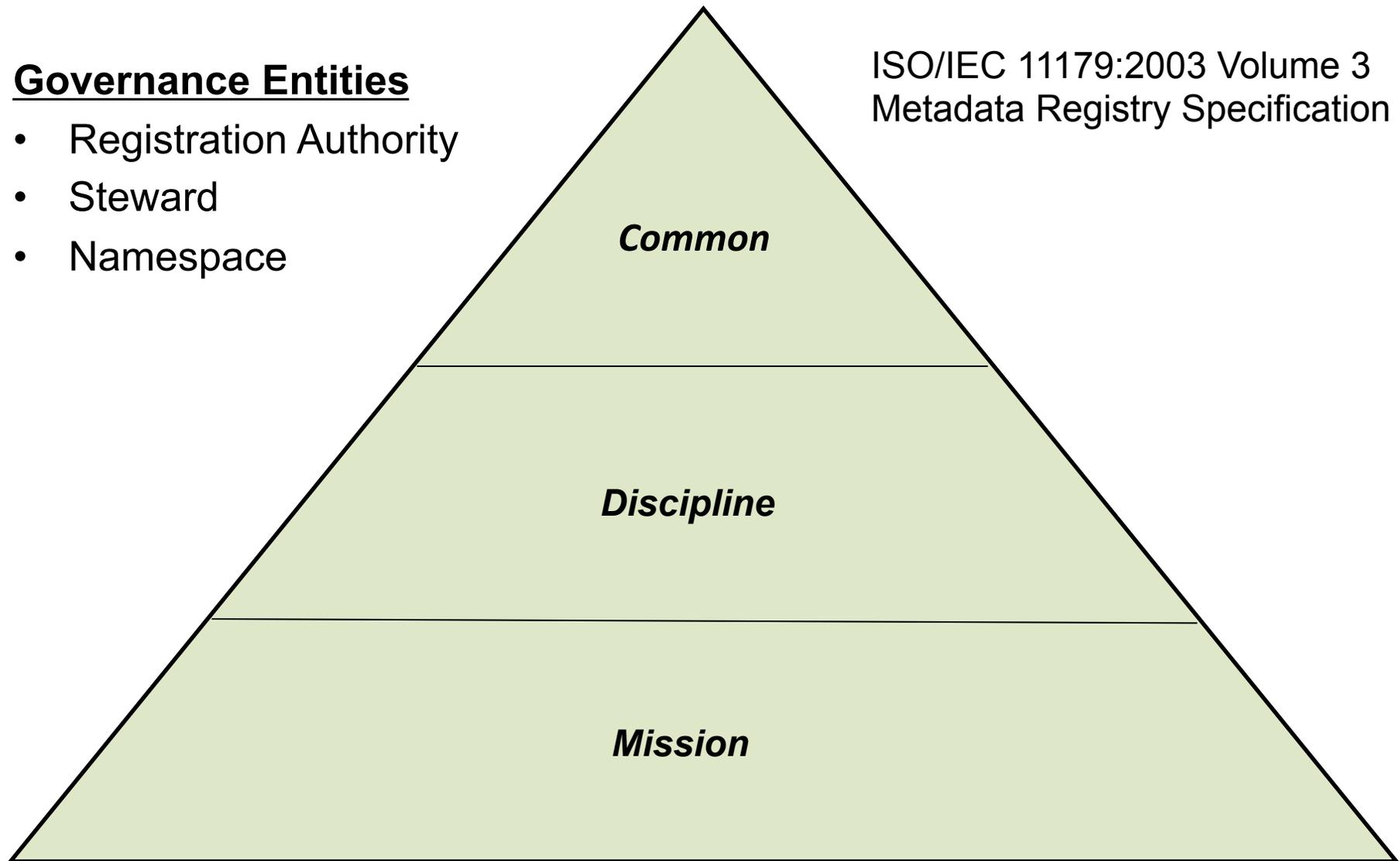
# Data Dictionary Governance

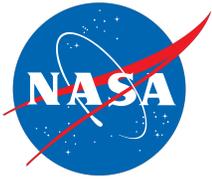


## Governance Entities

- Registration Authority
- Steward
- Namespace

ISO/IEC 11179:2003 Volume 3  
Metadata Registry Specification

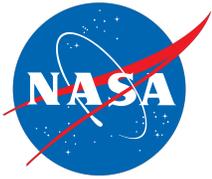




# Topics



- Introduction
- Key Requirements
- Build 3b Deliverables
- Change Requests for Version 1.0
- Conclusion



## Requirements For The PDS4 Data Standards



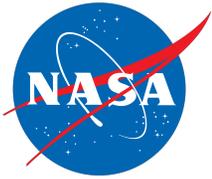
<b>REQUIREMENT</b>	<b>IMPLEMENTATION</b>
1.3.X – Provide Data Dictionary	<ul style="list-style-type: none"><li>• The PDS4 data dictionary database was developed and is compliant with the ISO/IEC 11179 specification.</li><li>• It is used to produce both data dictionary documents and data dictionary products for the registry and data dictionary service.</li></ul>
1.4.1 PDS will define a standard for organizing, formatting, and documenting planetary science data	<ul style="list-style-type: none"><li>• The PDS4 Information Model defines the archive organization, data formats, and product labeling standards.</li><li>• The PDS4 Standards Reference documents additional requirements.</li></ul>
1.4.2 PDS will maintain a dictionary of terms, values, and relationships for standardized description of planetary science data	<ul style="list-style-type: none"><li>• The PDS4 Data Dictionary defines the attributes, classes, and relationships for defining planetary science data.</li></ul>
1.4.3 PDS will define a standard grammar for describing planetary science data	<ul style="list-style-type: none"><li>• XML and XML Schema 1.1 have been adopted for the PDS4 implementation.</li></ul>



# Requirements For The PDS4 Data Standards



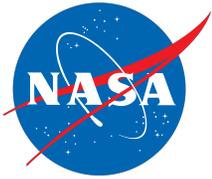
<b>REQUIREMENT</b>	<b>IMPLEMENTATION</b>
1.4.4 PDS will establish minimum content requirements for a data set (primary and ancillary data)	<ul style="list-style-type: none"><li>• The PDS4 Information Model defines observational and ancillary product types. These products are collected into PDS4 Collections and Archive Bundles.</li></ul>
1.4.5 PDS will, for each mission or other major data provider, produce a list of the minimum components required for archival data	<ul style="list-style-type: none"><li>• The PDS4 Information Model defines the archive bundle and its product collections. The archive bundle and its collections are customized for each mission.</li></ul>
3.1.2 PDS will develop and maintain online interfaces for discipline-specific searching	<ul style="list-style-type: none"><li>• The PDS4 Information Model and Data Dictionary defines information that is needed for search.</li></ul>
2.3.1 PDS will develop and publish procedures for determining syntactic and semantic compliance with its standards	<ul style="list-style-type: none"><li>• The adoption of XML and XML Schema 1.1 provide syntactic and semantic standards</li><li>• They provide utilities and tools for validation.</li></ul>



## MC Policy on PDS4 Data Formats

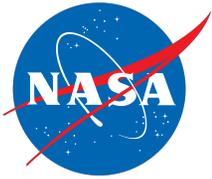


<b><i>Policy</i></b>	<b><i>Deliverable</i></b>
<p>PDS shall accept the following PDS4 data formats:</p>	
<ul style="list-style-type: none"><li>• Fixed-width binary and ASCII tables that are composed of identically structured records</li></ul>	<ul style="list-style-type: none"><li>• Table_Character and Table_Binary - Define heterogeneous repeating record of scalars.</li></ul>
<ul style="list-style-type: none"><li>• N-dimensional arrays of homogeneous binary elements</li></ul>	<ul style="list-style-type: none"><li>• Array - Defines a homogeneous N-dimensional array of scalars.</li></ul>
<ul style="list-style-type: none"><li>• Variable-width character 'spreadsheets' that are composed of repeating, multi-field, stream-delimited records.</li></ul>	<ul style="list-style-type: none"><li>• Delimited_Table - The Delimited_Table class defines a simple table (spreadsheet) with delimited fields and records.</li></ul>
<ul style="list-style-type: none"><li>• NAIF/SPICE files</li></ul>	<ul style="list-style-type: none"><li>• The SPICE_Kernel class describes SPICE files.</li></ul>
<ul style="list-style-type: none"><li>• ASCII text formats for PDS4 documentation. JPEG, GIF, and TIFF images for figures accompanying documents.</li></ul>	<ul style="list-style-type: none"><li>• Product_Document - A Product Document is a product consisting of a single logical document comprised of one or more document formats.</li><li>• ASCII Text is allowed as document formats.</li><li>• JPEG, GIF, TIFF, and PNG are allowed as non-science image formats.</li></ul>



# LADEE Product Types

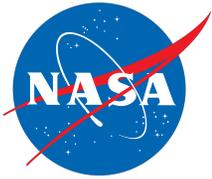
Instrument	Node	Data Format
LDEX	SBN	Table_Character
UVS	ATM	Table_Character
NMS	ATM	Table_Delimited
N/A	ATM	Product_Context (LADEE archive page)



# MAVEN Product Types

Instrument	Node	Data Format
SWEA	PPI	Array_1D*, 2D (PDS4 representation of CDF file)
SWIA	PPI	Array_1D*, 2D, 3D (PDS4 representation of CDF file)
STATIC	PPI	Array_1D*, 2D (PDS4 representation of CDF file)
SEP	PPI	Array_1D*, 2D (PDS4 representation of CDF file)
LPW	PPI	Array_1D*, 2D (PDS4 representation of CDF file)
LPW-EUV	PPI	Array_1D*, 2D (PDS4 representation of CDF file)
MAG	PPI	Table_Character
IUVS	ATM	Array_2D, Array_3D, Table_Binary, Table_Character
NGIMS	ATM	Table_Character
ACC	ATM	Table_Character
N/A	PPI	Product_Context (MAVEN archive page)

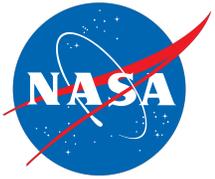
\* Array 1D support will be included in next release.



# Topics

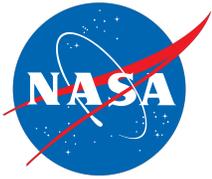


- Introduction
- Key Requirements
- Build 3b Deliverables
- Change Requests for Version 1.0
- Conclusion



# Build 3b

- The information model and standards were delivered as part of Build 3b (March 2013).
- Full testing occurred with Build 3b.
- Results presented to the MC in April 2013.
- MC voted to promote Build 3b to V1.0.0.0
  - It was later determined that two extensions to the common model were required to write MAVEN CDF products as compliant PDS4 files.



# Build 3b Deliverables\*



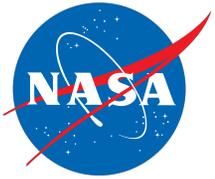
- Software System

- Registry Service
- Harvest Tool
- Validate Tool \*
- Security Service
- Report Service
- Search Service
- Catalog Ingest Tool
- Upgraded portal search and page views to support PDS4

- Data Standards\*

- Information Model
- XML Schemas
- Data Dictionary
- Concepts Document
- Standards Reference
- Data Providers Handbook
- PDS4 Example Products

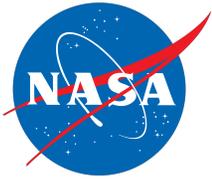
\* Posted to <http://pds.nasa.gov/pds4>



# Topics



- Introduction
- Key Requirements
- Build 3b Deliverables
- Change Requests for Version 1.0
- Conclusion



# CCB Changes to Support LADEE/MAVEN



<b><i>CCB Item</i></b>	<b><i>Deliverable</i></b>
<ul style="list-style-type: none"><li>• 1 Dimensional Array for MAVEN - CCB-17, Aug 20</li></ul>	<ul style="list-style-type: none"><li>• Array_1D - An “extension” of the PDS4 Array fundamental data structure.</li></ul>
<ul style="list-style-type: none"><li>• Data Format Header for MAVEN - CCB-24, Aug 20</li></ul>	<ul style="list-style-type: none"><li>• Header (CDF) – An additional permissible value for the Header attribute &lt; parsing_standard_identifier &gt;. Provides format identification to user/software.</li></ul>
<ul style="list-style-type: none"><li>• Partial redesign of the Primary_Result_Summary - CCB-36*, Sep 24</li></ul>	<ul style="list-style-type: none"><li>• Science_Facets – This new subclass enhances faceted searching across the entire PDS registry for observational and other science data.</li></ul>

\* CCB-36 is the result of a cross node WG. It has been reviewed by the DDWG, prototyped and is ready for CCB review and approval.

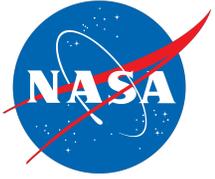




## Next IM Release

- Minor changes to Version 1.0.0.0 will enable support for LADEE and MAVEN and complete the data provider support
  - A simple “extension” to the array class for array 1d which is analogous to array 2d, 3d, etc.
  - A new permissible header value for CDF
  - An interim version of the model with these changes has been generated for testing which works with build 3b software (validate, harvest, register)
  - The current plan is to deploy with the next major build which fits the MAVEN schedule.
- An upgrade of the Primary Results Summary metadata will improve product level search for the disciplines.
  - This is planned to be deployed with the distribution services in build 4a.
  - PPI and Atmos nodes plan to adopt for LADEE/MAVEN

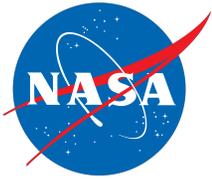




# Topics

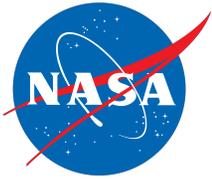


- Introduction
- Key Requirements
- Build 3b Deliverables
- Change Requests for Version 1.0
- Conclusion



# Conclusion

- The information model is very stable
  - Extensive testing and review by PDS, IPDA, missions, and external reviewers
- The LADEE and MAVEN planned products types have been tested against the IM
  - Minor changes for CDF have been easily accommodated
  - Product types have been validated, harvested and registered
- CCB in place to manage changes going forward



# Acknowledgements\*

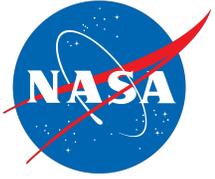


Ed Bell  
Richard Chen  
Dan Crichton  
Amy Culver  
Patty Garcia  
Ed Grayzeck  
Ed Guinness  
Mitch Gordon  
Sean Hardman  
Lyle Huber  
Steve Hughes  
Chris Isbell  
Steve Joy  
Ronald Joyner

Debra Kazden  
Todd King  
Joe Mafi  
Mike Martin  
Stephanie McLaughlin  
Thomas Morgan  
Lynn Neakrase  
Paul Ramirez  
Anne Raugh  
Mark Rose  
Elizabeth Rye  
Boris Semenov  
Dick Simpson  
Susie Slavney

David Heather  
Santa Martinez  
  
Peter Allan  
Michel Gangloff  
Thomas Roatsch  
Alain Sarkissian

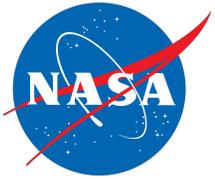
\* Anyone who sat through a DDWG 2-hour telecon or provided useful input.



# Thank You

# Questions and Answers





# Backup

