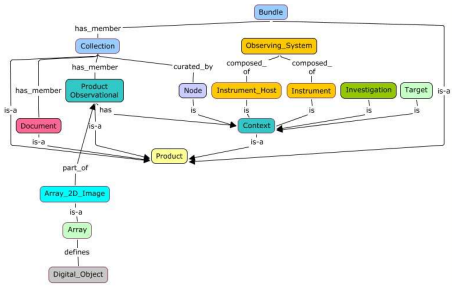
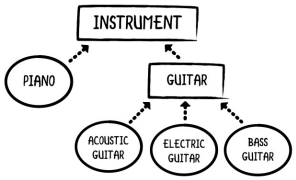
The PDS4 Information Model Knowledge Graph

John S. Hughes, Jordan Padams, Bess Schrader,
Ann Bernath, Thomas Loubrieu, Daria Topousis

PSIDA 2022
Planetary Science Informatics And Data Analytics Conference
ESAC, Madrid, Spain
21-23 June 2022





What is the problem?

- A data ecosystem is a collection of infrastructure, analytics, and applications used to capture and analyze data.
- A key goal of a data ecosystem is to provide the users with a consistent and unified view of the data and the metadata.
  - Intuitively a common “vocabulary” for the data and the metadata is necessary.
- **In actuality a common “vocabulary” is typically not available and in any case are very difficult to produce and maintain.**
  - Example: “Packet” has three different definitions in documents produced by the Consultative Committee for Space Data Systems (CCSDS). 1982

Planetary Data Ecosystem (PDE)

- NASA desires a seamlessly integrated Planetary Data Ecosystem [1]
  - The goal is to improve the planetary science community's access to, and use of, high-quality data
  - This intuitively requires a common "vocabulary" (knowledge base) of data and metadata terms, definitions, and relationships across many data sources and scientific domains.

  - *The planetary scientific community continues to change, evolving not only scientifically and technically but organizationally.*

[1] PDE IRB (2021) *Final Report of the Planetary Data Ecosystem Independent Review Board*.

Knowledge Graph

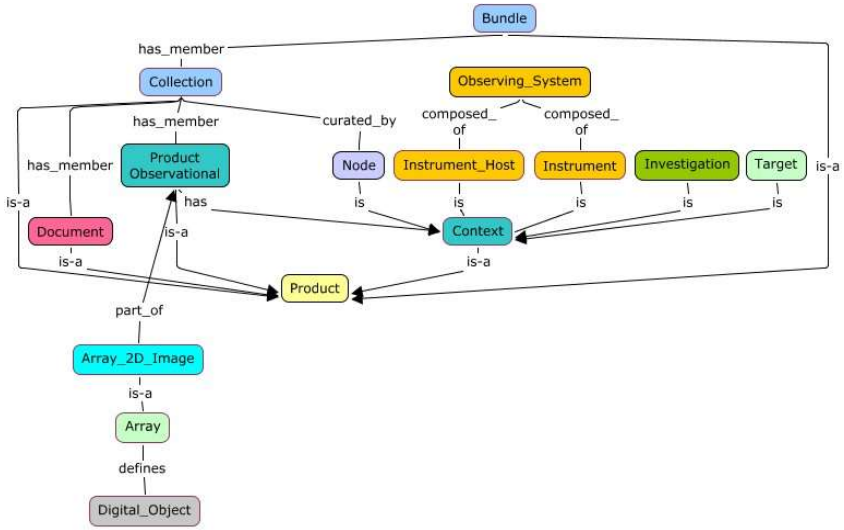
- “A knowledge graph stores interlinked descriptions of entities – objects, events, situations or abstract concepts – while also encoding the semantics underlying the used terminology.” [1]
  - No longer simply a research topic in Artificial intelligence

- A knowledge graph is a valuable tool for understanding specific domains
  - Provides a common context for the data and the metadata
  - Harmonizes (or at least relates) terms using common semantic models
    - Helps resolve inconsistencies and ambiguities
  - Improves operational efficiency and enables discovery
  - Can infer new knowledge

[1] “What is a Knowledge Graph?“, OntoText, 2018, https://www.ontotext.com/knowledgehub/fundamentals/what-is-a-knowledge-graph

PDS4 Information Model[1]

- Defines real-world entities and relationship in the planetary science domain
  - Designed for the purposes of long-term preservation and reuse of the data
- Contains sufficient information for constructing a knowledge graph.



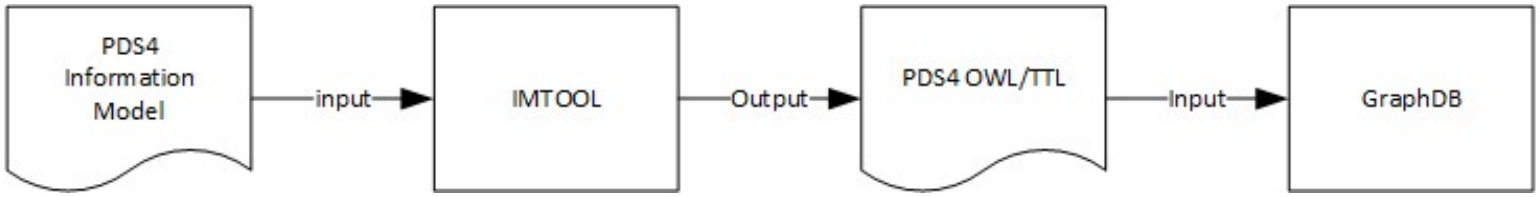
[1] Hughes, et.al, (2009) Ontology-Based Information Model Development for Science Information Reuse and Integration

Towards “Seamless” Integration

- A knowledge graph has been generated from the PDS4 Information Model
  - Phase 1 includes PDS4 Products and their component classes and relationships
  - Phase 2 will include class attributes.
- As a test case, the team has started to identify common concepts and entities across the PDS IM knowledge graph and JPL’s Institutional Knowledge Graph (IKG) [1]
  - The IKG is a centrally maintained knowledge graph identifying and describing JPL’s common concepts, such as people, organizations, facilities, and project.
  - The task is being performed by an expert team of knowledge engineers.

Generating the PDS4 Knowledge Graph

- The PDS4 Information Model is stored as a set of related ontologies.
- IMTool marshals the ontologies into memory and export the contents to the selected file format.



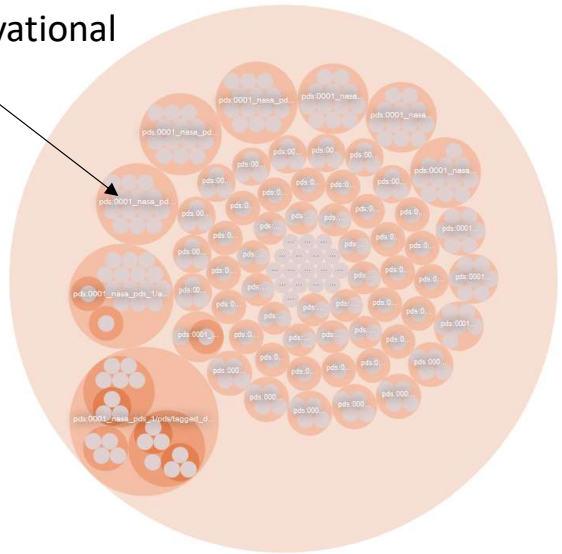
- W3C Web *Ontology* Language (*OWL*) is a Semantic Web language designed to represent rich and complex knowledge about things, groups of things.
  - TTL is a compact syntax alternative to RDF/XML; Resource Description Framework (RDF) is a standard for describing web resources and data interchange.
- The OWL/TTL file is loaded into GraphDB
  - Ontotext GraphDB is a graph database and knowledge discovery tool compliant with RDF and SPARQL

PDS4 IM -> OWL/TTL -> GraphDB

Product_Observational

Name: **Product_Observational**			***Version Id:*** **1.7.0.0**
Description: A Product_Observational is a set of one or more information objects produced by an observing system.			
Namespace Id: pds	*Steward:* pds	*Role:* concrete	*Status:* Active
Class Hierarchy: Product :: Product_Observational			
No Attributes			
Association(s)	*Name*	*Cardinality*	*Class*
	has_identification_area	1..1	Identification_Area
	observation_area	1..1	Observation_Area
	reference_list	0..1	Reference_List
	file_area	1..Unbounded	File_Area_Observational
	file_area_supplemental	0..Unbounded	File_Area_Observational_Supplemental
Referenced from: None			

Product Observational



- ### Class (Selected):0001_NASA_PDS_1.pds.Product_Observational ###
- <http://ontology.pds.nasa.gov/pds/0001_nasa_pds_1/pds/product/pds/product_observational-v1.0> rdf:type rdfs:Class;
- rdfs:label "Product_Observational";
- pds:subClassOf <http://ontology.pds.nasa.gov/pds/0001_nasa_pds_1/all/user/pds/product-v1.0>;
- rdfs:comment "A Product_Observational is a set of one or more information objects produced by an observing system.".

- ### ObjectProperty (Multi):0001_NASA_PDS_1.pds.Product_Observational ###
- <http://ontology.pds.nasa.gov/pds/0001_nasa_pds_1-pds-componentof-product_observational-v1.0> rdf:type owl:ObjectProperty;
- rdfs:label "components of Product_Observational";
- rdfs:domain <http://ontology.pds.nasa.gov/pds/0001_nasa_pds_1-product_observational-components-v1.0>;
- rdfs:range <http://ontology.pds.nasa.gov/pds/0001_nasa_pds_1/pds/product/pds/product_observational-v1.0>;
- rdfs:subPropertyOf <http://ontology.pds.nasa.gov/pds/urn/nasa/pds/0001_NASA_PDS_1/componentof-v1.0>;
- rdfs:comment "These are the component(s) of Product_Observational".

PDS Instrument Types (c 2016 - ~80)

ACCELEROMETER
ALPHA PARTICLE SPECTROMETER
ATM - ACCELERATION INSTRUMENT
ATM - ALTITUDE INSTRUMENT
ATM - COMPOSITION INSTRUMENT
ATM - DENSITY INSTRUMENT
ATM - HELIUM DETECTOR
ATM - NEPHELOMETER
ATM - PRESSURE INSTRUMENT
ATM - PYROLYSER
ATM - RAINFALL INSTRUMENT
ATM - TEMPERATURE INSTRUMENT
ATM - WIND INSTRUMENT
ATOMIC FORCE MICROSCOPE
DUST INSTRUMENT
GAMMA RAY SPECTROMETER
GAMMA-RAY SPECTROMETER
GRAVITY INSTRUMENT
HIGH ENERGY PARTICLE DETECTOR
H-SPECTRUM SPECTROMETER
INFRARED IMAGER
INFRARED INTERFEROMETER
INFRARED PHOTOMETER
INFRARED RADIOMETER
INFRARED SPECTROMETER
INFRARED SPECTROMETER FIX
LASER ALTIMETER
LIDAR IMAGER
LOW ENERGY PARTICLE DETECTOR

MAGNETOMETER
MASS SPECTROMETER
MICROWAVE PHOTOMETER
MICROWAVE RADIOMETER
NEAR INFRARED IMAGER
NEAR INFRARED RADIOMETER
NEAR INFRARED SPECTROMETER
NEUTRON DETECTOR
NEUTRON SPECTROMETER
PHOTOMETER
PLASMA WAVE SPECTROMETER
POLARIMETER
RADAR
RADIO SCIENCE SUBSYSTEM
RADIO TELESCOPE
RADIOMETER
RADIOMETER PACKAGE
SEISMOMETER
SHORTWAVE INFRARED IMAGER
SHORTWAVE INFRARED SPECTROMETER
SMALL BODIES SCIENCES
SOIL - CHEMISTRY INSTRUMENT
SOIL - CONDUCTIVITY INSTRUMENT
SOIL - EH INSTRUMENT
SOIL - ELECTRICAL CONDUCTIVITY INSTRUMENT
SOIL - GLOBAL POSITIONING SYSTEM
SOIL - GRIND ENERGY INSTRUMENT
SOIL - HUMIDITY INSTRUMENT
SOIL - HYGROMETER
SOIL - ISOTOPIC RATIOS INSTRUMENT

SOIL - METABOLISM INSTRUMENT
SOIL - MINEROLOGY INSTRUMENT
SOIL - ORGANIC MOLECULES INSTRUMENT
SOIL - ORGANICS INSTRUMENT
SOIL - PENETROMETER
SOIL - PH INSTRUMENT
SOIL - ROBOTIC ARM INSTRUMENT
SOIL - TEMPERATURE INSTRUMENT
SOIL - THERMAL CONDUCTIVITY INSTRUMENT
SOIL - THERMISTOR
SOIL - VOLATILES INSTRUMENT
SPECTROMETER - PACKAGE
SPECTROMETER PACKAGE
SYNTHETIC APERTURE RADAR
THERMAL IMAGER
THERMAL INFRARED SPECTROMETER
ULTRAVIOLET IMAGER
ULTRAVIOLET RADIOMETER
ULTRAVIOLET SPECTROGRAPH
ULTRAVIOLET SPECTROMETER
VISIBLE IMAGER
VISIBLE PHOTOMETER
VISIBLE RADIOMETER
VISIBLE SPECTROGRAPH

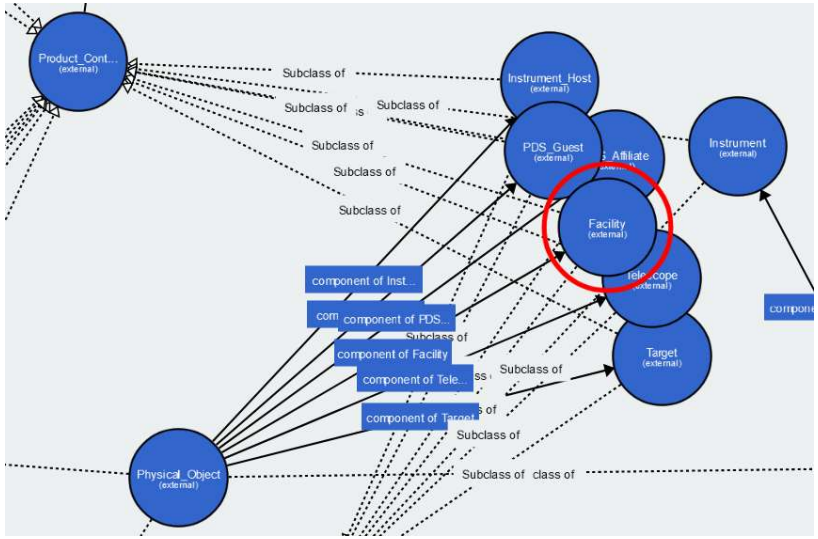
PDS4 Instrument Types (~40)

Accelerometer - An accelerometer measures acceleration -- rate of change of velocity -- in
Altimeter - An altimeter measures distance above a surface.
  - alt - Lidar
Atmospheric Structure Instrument - An atmospheric structure instrument measures one
  - alt - Meteorology
  - alt - Weather Station
Camera - A camera is an optical instrument that captures a still image or a sequence of
  - alt - Imager
Charged Particle Detector - A charged particle detector detects and/or counts charged part
Dust Analyzer - A dust analyzer measures the size and/or energy distribution of dust parti
Electric Field Instrument - An electric field instrument measures the direction and/or str
Gamma Ray Detector - A gamma ray detector is an instrument that detects gamma rays.
Gas Analyzer - A gas analyzer measures the concentration of one or more species in a m
Gravimeter - A gravimeter measures gravitational acceleration.
Imager - An imager detects and converts information into a digital image.
  - alt - Camera
Imaging Spectrometer - An imaging spectrometer acquires a spectrally-resolved image of an
Interferometer - An interferometer superposes waves such that constructive and destructive
Langmuir Probe - A Langmuir probe consists of one or more electrodes used to determine in-
Lidar - A lidar measures distance to a target by illuminating it with a pulsed laser
  - alt - Altimeter
Magnetometer - A magnetometer measures the direction and/or strength of a magnetic field.
Mass Spectrometer - A mass spectrometer sorts and counts atoms, ions, and/or molecules bas
Meteorology - A meteorology instrument measures in situ meteorological conditions.
  - alt - Weather Station -- exactMatch
  - alt - Atmospheric Structure Instrument
Microphone - A microphone converts sound waves into electrical signals.
Microscope - A microscope magnifies objects that are too small to be seen with the naked e
Mutual Impedance Probe - A mutual impedance probe measures in situ bulk plasma properties
Nephelometer - A nephelometer measures the concentration of suspended (cloud) particulates
Neutral Particle Detector - A neutral particle detector detects and/or counts neutral part
Neutron Detector - A neutron detector detects and/or counts neutrons.
Photometer - A photometer measures the intensity of electromagnetic radiation over a wavel
Polarimeter - A polarimeter measures the polarization of an electromagnetic wave.
Radar - A radar transmits an electromagnetic wave, then measures amplitude, time delay, fr
Radio Receiver - A radio receiver detects the information in propagating electromagnetic w
Radio Science - Radio science is the use of active and/or passive electromagnetic waves to
Radiometer - A radiometer measures radiant flux (power) of electromagnetic radiation.
Relaxation Sounder - A relaxation sounder determines the properties of a plasma
Seismometer - A seismometer measures ground motions such as might be caused by earthquakes
Spacecraft Sensor - A spacecraft sensor captures information on the status or physical con
Spectrometer - A spectrometer measures an energy spectrum.
  - alt - Spectrograph -- closeMatch
Spectrum Analyzer - A spectrum analyzer measures the properties of photons, charged partic
Sub-Surface Tool
Surface Tool
Temperature Sensor
Weather Station - A weather station is a suite of instruments that measures in situ meteor
  - alt - Meteorology -- exactMatch
  - alt - Atmospheric Structure Instrument
Wind Tunnel - A wind tunnel is used to study the effects of air moving past solid objects.

PDS4 IM -> OWL/TTL -> WebVOWL
Facility -- Context -- Physical

Facility

Name: **Facility**			***Version Id:*** **1.0.0.0**
Description: The Facility class provides a name and address for a terrestrial observatory or laboratory.			
Namespace Id: pds	*Steward:* pds	*Role:* concrete	*Status:* Active
Class Hierarchy: Tagged_NonDigital_Object :: TNDO_Context :: Facility			
Attribute(s)	*Name*	*Cardinality*	*Value*
	name	0..1	None
	type	1..1	Laboratory, Observatory
	address	0..1	None
	country	0..1	None
	description	0..1	None
Association(s)	*Name*	*Cardinality*	*Class*
	data_object	1..1	Physical_Object
Referenced from: Product_Context			



- ### http://ontology.pds.nasa.gov/pds/0001_nasa_pds_1/pds/tndo_context/pds/facility
- <http://ontology.pds.nasa.gov/pds/0001_nasa_pds_1/pds/tndo_context/pds/facility> rdf:type owl:Class ;
- rdfs:subClassOf <http://ontology.pds.nasa.gov/pds/0001_nasa_pds_1-product_context-components> ,
- <http://ontology.pds.nasa.gov/pds/0001_nasa_pds_1/pds/tagged_nondigital_object/pds/tndo_context> ;
- <http://ontology.pds.nasa.gov/pds/subClassOf> <http://ontology.pds.nasa.gov/pds/0001_nasa_pds_1/pds/tagged_nondigital_object/pds/tndo_context> ;
- rdfs:comment "The Facility class provides a name and address for a terrestrial observatory or laboratory." ;
- rdfs:label "Facility" ;
- owl:versionInfo "1.0.0.0" .

Definitions of "*Facility*"

facility, installation (a building or place that provides a particular service or is used for a particular industry) *"the assembly plant is an enormous facility"*

Source: Wordnet

Broader

"A Facility is a physical location provided for a particular purpose."

Source: JPL's' Institutional Knowledge Graph (IKG)

Narrower

The Facility class provides a name and address for a terrestrial observatory or laboratory."

Source: PDS4 Information Model

Narrower

Low Temperature Microgravity Physics Facility + 1000s of others

Source: NASA Technical Reports Server - Dictionary of Technical Terms for Aerospace Use

Summary

- The goal is to develop a Planetary Science Knowledge Graph, a centralized, machine-readable repository for metadata/data terms, meanings, and relationships
- Provides the “Context Information” necessary for the near and long-term reuse of scientific data.
- Used to optimize information collection, organization, search, and retrieval.
- Leverages existing semantic technologies, techniques, and expertise.
- Supports the data ecosystem by providing a gold-source for terms, definitions, and relationships.

Acknowledgements

The authors would like to thank the PDS Data Design Working Group (DDWG) and other PDS Discipline Node staff for their individual roles in the development of the PDS4 Information Model.

Backup