THE PDS4 INFORMATION MODEL KNOWLEDGE GRAPH

John S. Hughes[1], Jordan Padams [2], Bess Schrader [3], Ann Bernath[4], Thomas Loubrieu[5], Daria Topousis[6]
[1]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA
[1]John.S.Hughes@jpl.nasa.gov, [2]Jordan.H.Padams@jpl.nasa.gov, [3]Bess.P.Schrader@jpl.nasa.gov,
[4]Ann.M.Bernath@jpl.nasa.gov, [5]Thomas.G.Loubrieu@jpl.nasa.gov, [6]Daria.E.Topousis@jpl.nasa.gov

Introduction: A modern scientific data ecosystem consisting of a network of evolving digital repositories needs to combine data from different sources to provide users with a unified view of data. One approach to data integration relies on a knowledge graph that captures the interrelationships between the data items across the digital repositories. The mappings are created on as needed basis in response to addressing specific science questions. The science questions themselves can be represented within a knowledge graph.

Knowledge graphs: Knowledge graphs are no longer simply research topics in Artificial intelligence but are now valuable tools for understanding specific domains and harmonizing data into common semantic models to improve operational efficiency and enable discovery. Knowledge graphs accomplish this by defining the real-world entities in a domain, providing relationships between them, and inferring new classes of entities and relations. In the development process, extensions can be made to the knowledge base and inconsistencies and ambiguities are resolved.
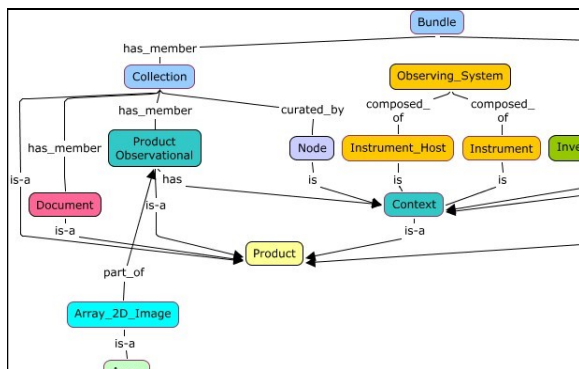

Figure 1: Concept Map

Information Model: The PDS4 Information Model [2] defines real-world entities and relationship in the planetary science domain for the purposes of long-term preservation and reuse of the data. The information model contains sufficient information for constructing a knowledge graph. Figure 1 depicts entities in the upper level of the model as a concept map and illustrates a few of the interrelationships.

However, science domains change and the planetary sciences continue to evolve not only scientifically and technically but organizationally. For example, NASA desires to develop a seamlessly integrated Planetary Data Ecosystem [1] that will improve the planetary science community's access to, and use of, high-quality data. This results in the need to harmonize data across several data sources.

Conclusion: A knowledge graph has been generated from the PDS4 Information Model to support such a task. This presentation will provide an overview of the technical issues involved in generating the knowledge graph. It will also describe the progress so far in integrating the knowledge graph with JPL's Institutional Knowledge Graph (IKG), a centrally maintained knowledge graph identifying and describing JPL's common concepts, such as people, organizations, facilities, and projects.

References:
[1] PDE IRB (2021) *Final Report of the Planetary Data Ecosystem Independent Review Board*.
[2] Hughes, et.al, (2009) *Ontology-Based Information Model Development for Science Information Reuse and Integration*.