Analysis Ready Data Available Through the SpatioTemporal Asset Catalog (STAC) Specification: Investigating the Application to Planetary Data. R. L. Fergason[1], M. A. Hunter[1], J. R. Laura[1], T. M. Hare[1]. [1]U.S. Geological Survey, Astrogeology Science Center, Flagstaff, AZ, USA, rfergason@usgs.gov.

Introduction: Most publicly available planetary data are not analysis-ready, meaning that individuals must download raw or reduced data from mission websites or the Planetary Data System (PDS), radiometrically and geometrically calibrate and photogrammetrically correct the data, often using multiple software packages before analysis can begin. This effort is significant and requires individuals to develop specialized processing and tool knowledge unrelated to their scientific expertise and for each dataset of interest. This current model is a barrier to using new planetary data sets, takes effort away from data analysis and science investigations, likely reduces the scientific output of the planetary science community, and potentially inhibits innovative science and collaboration. In addition, the burden is disproportionately felt by those new to a field or data set, including graduate students and individuals not connected to mission teams.

In contrast, the terrestrial community typically has access to analysis-ready data products (e.g., Landsat 8 [1] and Sentinal-2 [2]) that are ready for immediate use in a Geographical Information System (GIS) or other data analysis tools. This broad access to analysis-ready data significantly reduces redundant data processing efforts. We look to the terrestrial community to gain insight into these methodologies and then seek to adapt these techniques to meet the needs of the planetary community. By developing a means by which the planetary community could have ready-access to calibrated and geometrically projected data using well-documented and proven methods, the need for individuals to learn and complete these tasks would be reduced, providing additional resources for science investigations and potentially increasing the science return from R&A programs. In addition, the availability of controlled data products with improved image position knowledge could be significantly increased.

Petabytes of raw and derived planetary spatial data have been collected and generated throughout the planetary community, with the potential to greatly increase our understanding of the solar system and improve mission and policy decisions. Historically, tools have been provided (e.g., ISIS software system) to allow individual researchers to process data in a consistent, reliable, and repeatable way. As planetary data evolves and the volume of data collected continues to increase, we expect that there will be ever-increasing value to data pre-processed using standard and demonstrated methods. Furthermore, as data migrates to the Cloud we have an opportunity to make these data more accessible and interoperable. Our goal for the efforts described herein is to enable planetary data to be more easily discoverable in a format that is analysis ready. We envision such a service to both allow individuals to discover and utilize discrete images or derived products for scientific or engineering purposes and for existing services and portals to consume data via this service through a well-documented Application Programming Interface (API).

The STAC Specification: We are currently adapting, implementing, and testing one means by which the planetary community could have ready access to calibrated and geometrically projected data, discrete images and derived data products, using the Spatio-Temporal Asset Catalog (STAC; http://stacspec.org/; https://github.com/radiantearth/stac-spec) specification. The STAC specification provides a method to describe a range of geospatial metadata information, so it can more easily be indexed and discovered (c.f. http://stacspec.org/). The primary module for this service is a spatiotemporal asset, which is any file that represents information about a planetary body acquired at a specific location and time. This asset is typically the image or derived geospatial data product and we provide these data in the STAC-preferred Cloud Optimized GeoTiff format (COG). A JavaScript Object Notation (JSON) document is also created for the asset data, which is simple, straightforward, and can be customized for individual purposes (https://stacspec.org/ core.html). Through a planetary extension in development through efforts by this team, the core GeoJSON object and related structures can be modified to include information specific to the planetary science domain.

A primary element of this service is the STAC Item and represents a collection of related data and metadata (i.e., a collection of assets, as described above). A STAC Item is a GeoJSON object and can be easily read by any modern GIS or geospatial library. The STAC Item JSON specification includes fields for: 1) the time the asset represents; 2) a thumbnail for quick browsing; 3) asset links and links to the described data; and 4) relationship links which allow users to traverse other related STAC Items. A STAC Item can contain additional fields and JSON structures to enable data providers to expose additional metadata information and to enable the creation of additional tools. (c.f., https://stacspec.org/core.html, with modifications).

Results: We have investigated the feasibility, in terms of data standard requirements and software needs, for developing a method whereby radiometrically calibrated and geometrically projected data could be provided as a service to the planetary community using

STAC. Through this investigation, we have: 1) identified data products and standard formats we for initial testing; 2) developed the processing pipelines as a technology demonstration; and 3) defined preliminary middleware, backend, and frontend mechanisms for data delivery. We tested this implementation using the controlled Galileo and Voyager images of Europa [3] and the controlled Thermal Emission Imaging System (THEMIS) infrared data set of Mars [4]. We chose data from different missions to different planetary bodies to help ensure that our implementation was adaptable across different data sets and target bodies, and scalable between both smaller and larger planetary data sets. In addition, if the smithed kernels derived from a control network is not the default kernel in ISIS, then access to these controlled images requires significant data processing and kernel knowledge, is of high value to the planetary community, and these individual images are not currently available through another means.

We determined that the STAC specification can be applied to planetary data and our current progress includes understanding the STAC core specification and existing extensions, developing a prototype planetary extension, and generating a STAC compliant catalog of 125 Galileo image mosaics of Europa, 1204 controlled Galileo images of Europa, and of 29 controlled THEMIS infrared images of Mars.

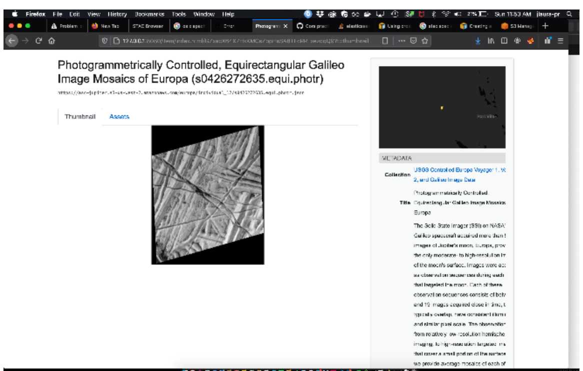

Figure 1. Example of STAC catalog browse page under development for Europa Galileo image mosaics, demonstrating the technical implementation and the typical product information available.

We have developed two prototype implementations to explore both the generation of a STAC catalog and how to provide data in a way that is discoverable and straightforward to ingest in a typical GIS. The Galileo and Voyager data of Europa are available in a basic catalog format where the user must currently search manually through the catalog to identify data products of interest (**Figure 1**). This implementation demonstrates a successful technical implementation of the STAC schema to these data. Although this is a necessary first step, we recognize that search methods are required to enable the discoverability of data of interest with greater ease and is the subject of current work efforts.

As data are available on the Cloud, the STAC schema provides the means to enable gaining access to geospatial data and analyzing those data without the requirement to download products locally. For the THEMIS data (**Figure 2**), we have developed a prototype QGIS plug-in that queries the STAC compliant catalog. Based on the user defining a location of interest, the COG images that fall within that user-defined location are then streamed directly to the QGIS application with no requirement to have the data available locally, although data can be downloaded and stored locally if desired (**Figure 2**). This plug-in eliminates the need for the user to download data onto their hard drive or a local server and then process these data before use. This QGIS plug-in is another technology demonstration that enabled us to ensure that our STAC catalogs were compliant and could be utilized by other standard tools.

Future work consists of further developing this means to deliver analysis-ready data to the community, including implementing state-of-the art search capabilities. In addition, if STAC continues to prove as promising as our initial investigation suggests, the leveraging of STAC standards in collaboration with the PDS will be a topic for future study.

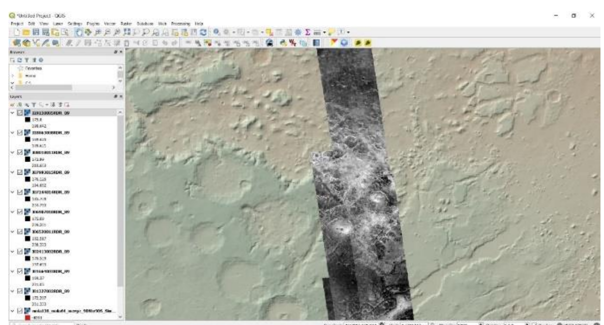

Figure 2. Prototype QGIS plug-in that queries a STAC compliant catalog of THEMIS infrared data of Mars and based on the user defining a location of interest. Cloud Optimized GeoTiff images were streamed directly to the application without downloading images locally.

Acknowledgements: References to commercial products are for identification purposes and do not imply an endorsement by the U.S. Government.

References: [1] Dwyer J. L. et al. (2018) *Remote Sens.*, doi:10.3390/rs10091363. [2] Drusch M. et al. (2012) *Remote Sens. Envi.*, doi:10.1016/ j.rse.2011.11.026. [3] Bland M. T. et al. (2020) *LPSC LI*, Abstract #1195. [4] Fergason R. L. and Weller L. (2019) *4th PDUW*, Abstract #7059.