INTEROPERABLE METHODS IN PLANETARY RESEARCH FOR GEOSPATIAL DATA ANALYSIS.

Trent M. Hare, Jason R. Laura, and Lisa R. Gaddis, Astrogeology Science Center, U.S. Geological Survey, 2255 North Gemini Drive, Flagstaff, AZ, 86001, (thare@usgs.gov).

Introduction: For more than a decade there has been a push in the planetary science community to support interoperable methods of accessing and working with geospatial data. Common geospatial data products for planetary research include image mosaics, digital elevation or terrain models, geologic maps, location databases (i.e., craters, volcanoes), or any data that can be tied to the surface of a planetary body (including moons, comets or asteroids).

Several U.S. and international research institutions have converged on geospatial standards such as: (1) standardized image formats which retain geospatial information (e.g., GeoTiff, GeoJpeg2000, PDS), (2) digital geologic mapping conventions, (3) planetary extensions for symbols that comply with U.S. Federal Geographic Data Committee cartographic and (4) geospatial metadata standards, and, notably, on-line mapping services as defined by the Open Geospatial Consortium (OGC).

Interoperable data formats and software interface standards are motivated by not only high-level product access, but also by the spatially distributed nature of big data stores and the need to support remote analytical interoperability. Herein, we provide an update of geospatial interoperability initiatives, and examples of their successful application.

Interoperable Tools: The U.S. Geological Survey's Astrogeology Science Center (ASC) is a major contributor of software for cartographic data processing for NASA missions and research programs, including the NASA's Planetary Cartography Program, Code S flight projects, research and data analysis projects, and the Planetary Data System (PDS). We support the Integrated Software for Imagers and Spectrometers (ISIS, [1]), a specialized image processing package for working with planetary image data. While it can ingest and export several different formats, it is really only able to process in its own specialized format (.cub). Goals for future versions of ISIS include the capability to process using several different formats, and to provide library access for development and scripting. In 2007, Geospatial Data Abstraction Library (GDAL) added reader support for the ISIS 3 format to improve interoperability. This reader is geared toward using products finalized by ISIS, not as a method to manipulate files within an ISIS workflow or pipeline.

GDAL: GDAL, released by the Open Source Geospatial Foundation (OSGeo), offers powerful capabilities for converting and processing planetary data. GDAL is a format translation library for geospatial raster and vector data [2]. In addition to the aforementioned ISIS3 reader, GDAL also supports ISIS2, PDS, and VICAR formats, allowing full interoperability with any GDAL supported application. Some popular applications with GDAL support include, QGIS, GRASS, MapServer, Esri's ArcMap and ArcGIS Pro, Generic Mapping Tools, and Opticks. For applications that do not use GDAL, the bundled routines released with GDAL can be used to convert these formats into more universal geospatial formats (e.g. GeoTiff).

Scripting Languages and GDAL: While GDAL is written in C/C++, it has bindings for use with many languages, including JAVA, PERL, Python, and .NET. As an interoperability example, we highlight Python which has a robust standard library and mature scientific computing stack (e.g. Numerical Python (NumPy), SciPy, Pandas, Matplotlib). GDAL provides the interface to support data reads into a common, in-memory format, the NumPy array. This opens a world of extremely powerful image processing methods. At ASC, we utilize Python for both rapid prototyping and production development. For example, to support the NASA InSight and Mars2020 missions, specialized topographic slope software was being supported in an outdated ISIS2 code base. Using GDAL, Python and existing array filtering functions in SciPy, we were able to quickly port the original source code, and integrate it with our digital terrain model workflow. During the port, we easily incorporated histogram binning (NumPy), to combine histogram and cumulative slope graphs (Matplotlib), and create colorized slope figures.

Interoperable Formats: By supporting GDAL, the need to standardize on a single format is greatly reduced. It has allowed us to more easily collaborate across different groups who prefer working with specific formats, either due to their preference or software requirements.

GeoTiff: The most popular geospatial format is GeoTiff [3]. The GeoTiff format, fully within the public domain, was created by Dr. Niles Ritter in the 1990's while working at Jet Propulsion Laboratory. GeoTiff makes use of a geospatial (cartographic) tags embedded within the TIFF file format. It is one of the only image formats which allows the flexibility to support tag structures without causing issues for applications that don't support those tags. The image format can support 8-bit grayscale images, and up to 16, 32 and 64 bit floating point elevation models. TIFF also supports a variety of compression and tiling options to increase the efficiency of image reading and online distribution. The BigTiff extension now allows single images to be greater than 4 gigabytes in size.

GeoJpeg2000: In recent years, the PDS has approved the use of the JPEG2000 format. This format supports the exact same tags as the GeoTiff format but stored within a Universally Unique Identifier (UUID) container. When utilized, this format is informally called GeoJPEG2000 (also GeoJP2™). In 2008, the HiRISE team was the first mission to release their map-projected PDS archives using a hybrid method which combines the use of the GeoJPEG2000 standard and a detached PDS label [4,5]. The simple text PDS label is necessary to hold required PDS metadata like author, instrument particulars, or mission dates which are not suitable for the geospatial container.

While this hybrid approach sounds like the best of both worlds, the JPEG2000 format does not yet support 32-bit floating point values, although it is part of the specification. There are PDS products that may simply not work well in this format. It is unfortunate that the open Jpeg2000 libraries (e.g. OpenJPEG or Jasper) are still far behind in capabilities and speed to proprietary solutions like the Kakadu library.

PDS4: Another upcoming format is the XML-based PDS4 [6]. This format was created to try to address some of the more confusing aspects of the PDS3 format. Like PDS3, PDS4 has a set of tags to support geospatial applications. Also, there is an initiative to allow the FITS format to include similar geospatial keywords [7]. Eventually, both PDS4 and "geo"FITS will be included for reading and potentially writing within the GDAL library.

Interoperable Web Services: The OGC is a consortium of more than 500 international companies, universities and government agencies which define standards such as the OGC Web Mapping Services (simple image maps), Web Feature Services (feature streaming), Web Coverage Services (rich scientific data streaming), and Catalog Services for the Web (data searching and discoverability). While these standards were developed for application to Earth-based data, they have been modified to support the planetary domain [8].

The strength of using an OGC approach is that it provides consistent data interfaces. OGC data-streaming standards are well-supported by enterprise level GIS and remote sensing applications, as well as light-weight and open-source data viewers.

Mapping Interoperability: The International Astronomical Union (IAU) defines the recommended rotation rate, spin axis, prime meridian, and reference surface for planets and satellites; however, their oversight does not cover other standards essential for digital mapping including common feature attributions, feature symbols, recommended mapping scales and finally the documentation of the data. When possible, it is recommended that digital maps use these standards so that consistent map products can be developed.

Feature attributes and their assigned symbols for planetary digital maps are commonly defined in the Digital Cartographic Standard for Geologic Map Symbolization [9] prepared by the USGS for the Federal Geographic Data Committee (FGDC). For example, recommended attributes for geologic contacts or geologic unit boundaries include attributes for contact certainty. The nominal level categories are then given explicit symbolic representations such as solid black line for certain contacts or dashed black lines for approximate boundaries. Symbology is primarily drawn from the same set of attributes and symbols as used for Earth. This heritage facilitates the understanding of geologic or thematic planetary maps because readers are familiar with the feature attribution names and symbol types. This common vocabulary supports interoperability [10].

Standardized Metadata: In short, metadata is the ancillary documentation that helps describe the rationale, authorship, attribute descriptions, spatial reference, and other pertinent information for data. For planetary data, PDS archives are the recommended method to document data products. Unfortunately, PDS metadata is not readily supported in more widely used geospatial data portals. Most geospatial portals require metadata as defined by the FGDC or International Organization for Standardization (ISO). Methods for conversion from PDS to FGDC/ISO metadata standards should be possible given that the FGDC metadata standards only require a few minor additions to properly support planetary data [11] although we are targeting data sets that can be registered to a solid body (exempting, for the time being, products that focus on atmospheres, plasma, and rings).

Interoperable Data Portals: One of the latest trends in the geospatial community, including the planetary community, is to provide data portals. These portals help assemble data holdings for on-line browsing and download. Many of the terrestrial data portals are already built around the use of FGDC metadata, to import, describe, and catalogue the data, and OGC standards, for data access and transfer. Planetary data portals [12] could benefit from the lessons learned from terrestrial portals by offering robust data search capability leveraging these well-tested and implemented standards.

References: [1] Keszthelyi, L., et al. (2014), LPSC XLV, abs. #2546. [2] Hare, T.M., et al. (2007), LPSC XXXIX, abs #2536. [3] http://geotiff.osgeo.org [4] McEwen, A. S., et al. (2007), *J. Geophys. Res.*, 112, E05S02, doi:10.1029/2005JE002605. [5] Castalia, B., (2008), LPSC XXXIX, abs. #2484. [6] https://pds.nasa.gov/pds4 [7] Marmo, H., et al. (2016), *this volume*. [8] Hare, T., et al. (2001), LPSC XLI, abs. #2638. [9] Skinner, J., et al. (2016), Planetary Geologic Mapping Handbook., URL: http://planetarymapping.wr.usgs.gov/Page/view/Guidelines [10] Nass, A., et al. (2010), *Planet. Space Sci.*, doi:10.1016/j.pss.2010.08.022. [11] Hare, T.M., et al. (2011), LPSC 42, #2154. [12] Hare, T.M., et al. (2011), LPSC XLVI, #2476.