**GEOTIFF SUPPORT WITHIN A PDS4 ARCHIVE.** T.M. Hare and L.R. Gaddis, U. S. Geological Survey, Astrogeology Science Center, Flagstaff, AZ, thare@usgs.gov.

**Introduction:** The NASA Planetary Data System (PDS) has released version 4 of their required archival format, called PDS4 [1]. Since 2011, compliance with the PDS4 archiving standard has been required for data archives from NASA-funded missions and research investigations. The most noticeable change for users is the requirement for additional metadata and the change to the eXtensible Markup Language (XML) format. A major goal for the move to PDS4 is to improve discoverability by strongly enforcing consistency in metadata and allowing for explicit cross-referencing of data products.

Here we introduce the GeoTIFF format as an accepted "operational" PDS4 image format ([2], p. 12), where operational is defined as a primary science data product, for example raw, calibrated or derived image. Although, the use of GeoTIFF in PDS archives is best suited for derived or science-ready map projected image products.

**GeoTIFF:** The GeoTIFF format was originally created in the late 1990s by Dr. Niles Ritter while working at the NASA Jet Propulsion Laboratory and it was designed specifically to be used for derived geospatial image data. GeoTIFF is a public domain standard which allows geospatial information to be embedded within the Tagged Image File Format (TIFF). Such information or metadata allows the user to establish the spatial reference (or geolocation) within the file, including the map projection, coordinate system, body size definition, and other such parameters.

As almost all mapping and image processing packages support GeoTIFF, it has emerged as a standard image file format worldwide [3]. So much that, as of February 2017, GeoTIFF was recognized by the Library of Congress as the preferred archival format for cartographic raster images [4]. In September 2019, GeoTIFF version 1.1 was adopted as an Open Geospatial Consortium (OGC) Standard [5]. Under this GeoTIFF version 1.1, the OGC also formalized the existing GeoTIFF 1.0 specification by integrating it into OGC's standardization process. The TIFF format has been integrated into several International Organization for Standardization (ISO) standards (e.g. ISO 12234-2, ISO 12639). The GeoTIFF format is compatible with TIFF 6.0, so that applications that cannot read advanced geospatial metadata within GeoTIFF tags can still open the file.

**Technical Overview of TIFF File Structure:** The first version of the TIFF format was released in 1986 by Aldus Corporation for use in desktop publishing. Aldus published TIFF 6.0 in 1992, which was subsequently updated with an Adobe Systems copyright after they acquired Aldus in 1994. TIFF is a flexible file format for handling images with an associated descriptive header within a single file. While TIFF supports several compression methods and image tiling, ***it also supports the use of simple raw binary streams for image data, allowing TIFF to be fully compliant with PDS4 requirements.*** And although Adobe owns the copyright, there is no need for a license to implement software for reading and writing the TIFF format. A widely used TIFF library called *libtiff* was written and initially released in 1988 by Sam Leffler while working for Silicon Graphics. All *libtiff* source code is either in the public domain or under a permissive BSD-like license and it is free to use for any purpose (including within commercial software). *Libtiff* is now chiefly maintained by the Open Source Geospatial Foundation (OSGeo) and it is freely available, but we stress that any reader that supports reading a raw binary stream can also read such a PDS4-compliant GeoTIFF.

The TIFF format has a three-level hierarchy including a (1) Image File Header (IFH), (2) one or more directories called Image File Directories (IFDs) containing codes and their data, or a pointer to the data, and (3) the image data. The TIFF file always begins with an 8-byte header, which gives basic information about the file such as byte order (Little Endian or Big Endian), TIFF file ID or Version Number and a pointer to the first IFD. Traversing the IFD(s) allows one to locate the pointer (offset bytes) to the image data. Each IFD contains one or more data structures called tags. Each tag is a 12-byte record that contains a specific piece of information. The TIFF file format uses 32-bit offsets which limits the output file size to 4 gigabytes. Fortunately, BigTIFF, a simple variant of the original TIFF, uses 64-bit offsets and allows for much larger file sizes.

To support a geospatial (Geo/GIS) extension within the GeoTIFF label, an extra set of IFDs have been defined and standardized. The GeoTIFF format fully complies with the TIFF 6.0 specifications, and its extensions do not in any way contradict the TIFF standard, nor do they limit the scope of raster data supported by TIFF.

**Discussion:** Use of the TIFF/GeoTIFF format will help the planetary community and PDS to realize both a PDS4 compliant archival format and a highly usable image format supported by hundreds of graphics-based and scientific applications (including popular applications like Photoshop, MATLAB, ENVI, IDL, many GIS applications and common image libraries). Currently, many instrument teams and PDS Discipline Nodes already provide image data as GeoTIFF within "Extras" directories as part of their PDS3 archives, largely to enhance usability of the products. Thus, acceptance of PDS4-compliant TIFF or GeoTIFF file formats for future PDS archives could help cut archive data volumes significantly simply by eliminating duplicate copies.

**Converting to a PDS4-compliant GeoTIFF:** The conversion from dozens of common image formats, including many planetary formats (e.g. ISIS3 cubes, VICAR, FITS, PDS3) to a PDS4-compliant format is simple and can be accomplished using the open source library Geospatial Data Abstraction Library (GDAL). The GDAL PDS4 driver supports writing PDS4-compliant image data [7] as stored in a separate uncompressed raw file or an uncompressed GeoTIFF file (using a strip organization compatible with a raw file).

The GDAL routine that facilities conversions, *gdal_translate*, works on all major platforms (Windows, Macintosh, and Linux). Image parameters like the number of lines, samples, bands, bit type, machine type, and display direction will be automatically transferred to the output PDS4 label. For PDS4 label entries that cannot be automated by the GDAL driver (e.g., mission name), a PDS4 XML template file with this additional metadata can be used during conversion. This allows the data provider to use existing PDS4 tools like the PDS Label Assistant for Interactive Design (PLAID) and On-Line Archiving Facility (OLAF) to create the initial PDS4 template. Also, user-defined variables within the provided template label are supported and can be set during conversion. Lastly, the target body and radius, and when applicable, the map projection (e.g., Equirectangular), will be written to the Cartography section of the PDS4 label if defined in the source image. For *gdal_translate* conversion examples, see https://github.com/USGS-Astrogeology/GDAL_scripts/wiki.

**Summary:** The PDS now recognizes the TIFF/GeoTIFF format as an allowed PDS4 "operational" data format. For PDS4 compliance only the simplest TIFF format, a raw binary stream for the image data, is allowed. The image data as stored in such a TIFF/GeoTIFF can easily be described in the PDS4 label (e.g., lines, samples, bands, skip bytes, bit type, display direction. etc.). Highlights for allowing GeoTIFF in PDS include:

- GeoTIFF helps the PDS realize both an archival and an immediately usable format for the planetary science community. This is comparable to allowing the Flexible Image Transport System (FITS) file format used heavily within the astronomy community or the Common Data Format (CDF) used in the atmospheres community, both PDS4 formats.
- While TIFF supports compression and tiling, the only form that would be allowed in a PDS4 archive would be a simple raw binary stream (easily accessible via the defined skip bytes as defined in the PDS4 label).
- GeoTIFF is recognized as an archival format by the Library of Congress [4] and supported by the Open Geospatial Consortium (OGC), an internationally supported standards group [5]. Within the OGC supported 1.1 specification, there is an included lunar example, demonstrating its flexibility for the planetary domain.
- Since 2017, the open source GDAL library [6, 7] has included a robust implementation allowing for the creation of a PDS4-compliant TIFF/GeoTIFF, including checks to validate existing files.

**Acknowledgments:** This effort was supported by NASA's PDS Cartography and Imaging Sciences Node and Planetary Spatial Data Infrastructure (PSDI) InterAgency Agreement. Any use of trade, firm, or product names is for descriptive purposes and does not imply endorsement by the U.S. Government.

**References:** [1] PDS4 Information at the Planetary Data System, https://pds.nasa.gov/pds4/ [2] The PDS4 Data Provider's Handbook, 2019, v1.12.0, https://pds.nasa.gov/datastandards/documents/dph/v1/PDS4_DataProvidersHandbook_1.12.0.pdf [3] Hare, T.M. et al., 2018, Interoperability in Planetary Research for Geospatial Data Analysis, PSS, vol. 150, https://doi.org/10.1016/j.pss.2017.04.004.[4] GeoTIFF, Revision 1.0 , 2017, Sustainability of Digital Formats: Planning for Library of Congress Collections, https://www.loc.gov/preservation/digital/formats/fdd/fdd000279.shtml [5] OGC GeoTIFF Standard, 2019, https://www.opengeospatial.org/standards/geotiff [6] Hare, T.M. et al., 2018, Initial Support for PDS4 within the Geospatial Data Abstraction Library (GDAL), 49th Lunar and Planetary Science Conference, Abs. 1703, https://www.hou.usra.edu/meetings/lpsc2018/pdf/1703.pdf [7] GDAL documentation, Raster drivers, PDS4 – NASA Planetary Data System (Version 4), https://gdal.org/drivers/raster/pds4.html.