

Phoenix TEGA Bundle: Notes and Errata Regarding PDS3-to-PDS4 Migration

PDS Geosciences Node

April 17, 2026

Introduction

The Mars Phoenix Lander datasets were originally archived under NASA's Planetary Data System (PDS) PDS3 standard, and the PDS Nodes hosting the archives are actively migrating these archives to the PDS4 standard. Hereinafter, these notes apply specifically to the Phoenix Thermal Evolved Gas Analyzer (TEGA) datasets archived at the PDS Geosciences Node.

All data products in the original TEGA archive were PDS4-compliant, so no data reformatting was necessary. In the data directories, the metadata in PDS3 labels was copied to PDS4 labels, so that each data product has both a PDS3 and a PDS4 label. The PDS3 archive volume remains intact; no files have been removed. The only change is the addition of PDS4 labels and documentation files.

Archive Organization

Instead of the datasets and archive volumes in PDS3, products in PDS4 are organized into collections and bundles. A collection is a set of related products, which may be data products, document products, browse products, miscellaneous products, etc. A bundle is a set of related collections.

One PDS4 bundle has been defined for the TEGA dataset. The bundle has 11 data collections (one for each data type), two document collections, one calibration collection, and one miscellaneous collection. The table below shows the correspondence between the PDS3 volume and PDS4 bundle and PDS3 product types and PDS4 collections, respectively.

PDS4 Bundle and Corresponding PDS3 Volume

Archive	PDS4 Bundle ID	PDS3 Volume
Phoenix Thermal Evolved Gas Analyzer (TEGA) Data Bundle	urn:nasa:pds:phx_tega	PHXTEG_0001

PDS4 Collections and Corresponding PDS3 Datasets

PDS4 Collection	PDS4 Collection ID	PDS3 Data Set ID
Thermal scanning calorimetry raw data	data_scedr	PHX-M-TEGA-2-SCEDR-V1.0
Thermal scanning calorimetry processed data	data_scrdr	PHX-M-TEGA-4-SCRDR-V1.0
Mass Spectrometer sweep raw data	data_egaedr	PHX-M-TEGA-2-EGAEDR-V1.0
Mass Spectrometer sweep processed data	data_egsrdr	PHX-M-TEGA-4-EGSRDR-V1.0
Mass Spectrometer peak hopping raw data	data_eghedr	PHX-M-TEGA-2-EGHEDR-V1.0
Mass Spectrometer peak hopping processed data	data_eghrdr	PHX-M-TEGA-4-EGHRDR-V1.0

LED Sensor data	data_lededr	PHX-M-TEGA-2-LEDEDR-V1.0
TA and EGA engineering raw data	data_engedr	PHX-M-TEGA-2-ENGEDR-V1.0
TA and EGA engineering processed data	data_engrdr	PHX-M-TEGA-3-ENGRDR-V1.0
TEGA message log	data_msgedr	PHX-M-TEGA-2-MSGEDR-V1.0
Experimenter's Notebook	data_ekrdr	n/a

A bundle is identified by a file named bundle_*.xml (where * is some character sequence) in the root directory of a volume; this PDS4 label describes the bundle and lists the collections that belong to it. A collection, in turn, is identified by a PDS4 label named collection_*.xml in a subdirectory. The file collection_*_inventory.csv is found in the same directory as the collection label and lists the products that belong to that collection. Each product also has its own PDS4 label named *.xml. The PDS4 label and the data file it describes have the same base file name with a different extension.

Every bundle, collection, and product in PDS4 has a Logical Identifier (LID) that is guaranteed to be unique throughout the PDS. The LID is defined in the PDS4 label using the <logical_identifier> tag. For data products, the LID is analogous to PRODUCT_ID in a PDS3 label.

Not every PDS3 directory has a PDS4 counterpart. The PDS3 2008 (data), calibration, catalog, document, and index directories have been made into collections, but the extras and label directories have been left unchanged. No PDS4 labels have been provided for **aareadme.txt**, **voldesc.cat**, and **errata.txt** in the volume root directories. Note that the SIS documents in the document directory refer to the PDS3 datasets.

Notes

Errors in PDS3 labels for several data types were corrected in the PDS4 labels. No attempt was made to correct these errors in the PDS3 labels.

1. EGHEDR: REPETITIONS is incorrect.
2. EGHRDR: RECORDS is incorrect.
3. ENGRDR: Sol 147, activity 2, PRODUCT_ID and OPS_TOKEN are incorrect.
4. LEDEDR: REPETITIONS is incorrect.
5. SCEDR: STARTBYTE incorrect starting at column 7.
6. SCRDR: TEGA_TIME BYTES is incorrect.