Mars 2020 Project

Software Interface Specification (SIS)

PDS Camera Archive Bundle Structure

Version 2.0

Release Date: November 1, 2021

JPL Doc ID: D-105238

Prepared by: Robert Deen, Kate Crombie

Custodian: Robert Deen

Mission Reference: Mars 2020 Mission

Jet Propulsion Laboratory
California Institute of Technology

The research described in this document was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

TABLE OF CONTENTS

LIST OF TABLES

DOCUMENT CHANGE LOG

Version	Change	Date	Affected portion
0.1	Initial draft	2020-08-30	All
0.2	EDR peer review liens	2020-12-15	All
0.3	Cleanup for RDR Peer Review	2020-12-15	All
1.0	Final release 1 version	2021-08-01	All
2.0	Changed bundle structure for release 2	2021-09-29	All

TBD ITEMS

Item	Section(s)	Page(s)
All yellow background text in this document	All	

ACRONYMS AND ABREVIATIONS

Acronym	Meaning
ACI	Autofocusing Context Imager (SHERLOC camera)
ASCII	American Standard Code for Information Interchange
ATM	Atmospheres PDS node
EDR	Experiment Data Record
FDR	Fundamental Data Record
GEO	Geosciences PDS node
HTML	Hypertext Markup Language
IDS	Instrument Data System (at JPL)
IM	Information Model
IMG	Cartography and Imaging Science PDS node
JPL	Jet Propulsion Laboratory
LID	Logical Identifier
LIDVID	Versioned Logical Identifer
MCC	Micro-Context Camera (PIXL camera)
MEDA	Mars Environmental Dynamics Analyzer
MIPL	Multimission Instrument Processing Laboratory
MSL	Mars Science Laboratory
MSSS	Malin Space Science Systems
NASA	National Aeronautics and Space Administration
ODL	Object Description Language
OPGS	Operational Product Generation Subsystem
PDF/A	PDF for Archive
PDF	Portable Document Format

PDS	Planetary Data System
PDS4	Planetary Data System Version 4
PNG	Portable Network Graphics
PIXL	Planetry Instrument for X-ray Lithochemistry (Mars 2020 instrument)
RDR	Reduced Data Record
RTE	Return to Earth (Helicopter camera)
SOAS	Science Operations Analysis Subsystem
SHERLOC	Scanning Habitable Environments with Raman & Luminescence for Organics & Chemicals (Mars 2020 instrument)
SIS	Software Interface Specification
SOL	Mars Solar Day
SuperCam	(name) Mars 2020 instrument
TBD	To Be Determined/Delivered
TDR	Tile Data Record
u:n:p	Shorthand for urn:nasa:pds, the prefix for bundle names
URN	Uniform Resource Name
VICAR	Video Image Communication And Retrieval
VID	Version Identifier
Watson	(name) SHERLOC camera
XML	eXtensible Markup Language

GLOSSARY

Many of these definitions are taken from Appendix A of the PDS4 Concepts Document, https://pds.nasa.gov/datastandards/documents/concepts. The reader is referred to that document for more information.

Archive – A place in which public records or historical documents are preserved; also the material preserved. The term may also refer to all of PDS holdings – the PDS Archive.

Attribute – A property or characteristic that provides a unit of information. For example, 'color' and 'length' are possible attributes.

Basic Product – The simplest product in PDS4; one or more data objects (and their description objects), which constitute (typically) a single observation, document, etc. The only PDS4 products that are *not* basic products are collection and bundle products.

Bundle Product – A list of related collections. For example, a bundle could list a collection of raw data obtained by an instrument during its mission lifetime, a collection of the calibration products associated with the instrument, and a collection of all documentation relevant to the first two collections.

Class – The set of attributes (including a name and identifier) which describes an item defined in the PDS Information Model. A class is generic – a template from which individual items may be constructed.

Collection Product – A list of closely related basic products of a single type (e.g. observational data, browse, documents, etc.). A collection is itself a product (because it is simply a list, with its label), but it is not a *basic* product.

Data Object – A generic term for an object that is described by a description object. Data objects include both digital and non-digital objects.

Description Object – An object that describes another object. As appropriate, it will have structural and descriptive components. In PDS4 a 'description object' is a digital object – a string of bits with a predefined structure.

Digital Object – An object which consists of real electronically stored (digital) data.

Identifier – A unique character string by which a product, object, or other entity may be identified and located. Identifiers can be global, in which case they are unique across all of PDS (and its federation partners). A local identifier must be unique within a label.

Label – The aggregation of one or more description objects such that the aggregation describes a single PDS product. In the PDS4 implementation, labels are constructed using XML.

Logical Identifier (LID) – An identifier which identifies the set of all versions of a product.

Versioned Logical Identifier (LIDVID) – The concatenation of a logical identifier with a version identifier, providing a unique identifier for each version of a product.

Manifest - A list of contents.

Meshes – Also known as terrain models, meshes are high level products containing geometric description of the surface consisting of triangles, with image texture attached to each triangle.

Metadata – Data about data – for example, a 'description object' contains information (metadata) about an 'object.'

Object – A single instance of a class defined in the PDS Information Model.

ODL Label – A text-format label describing the data. ODL is the format used for PDS3; the distinction is that ODL defines the syntax of the label while PDS3 governs the content. All PDS3 labels are ODL but the converse is not true.

PDS Information Model – The set of rules governing the structure and content of PDS metadata. While the Information Model (IM) has been implemented in XML for PDS4, the model itself is implementation independent.

PDS Node – One of several organizations within PDS; each is responsible for archiving data within a given discipline or area of expertise.

Product – One or more tagged objects (digital, non-digital, or both) grouped together and having a single PDS-unique identifier. In the PDS4 implementation, the descriptions are combined into a single XML label. Although it may be possible to locate individual objects within PDS (and to find specific bit strings within digital objects), PDS4 defines 'products' to be the smallest granular unit of addressable data within its complete holdings.

Tagged Object – An entity categorized by the PDS IM, and described by a PDS label.

Registry – A data base that provides services for sharing content and metadata.

Repository – A place, room, or container where something is deposited or stored (often for safety).

Version Identifier (VID) – Consist of major and minor components separated by a "." (M.n), and identify a specific version of a product.

VICAR – Image processing system created and maintained at the JPL Multimission Image Processing Lab (MIPL). VICAR is used to create most of the image products in this archive. VICAR is available open source; see https://www-mipl.jpl.nasa.gov/vicar_open.html .

XML – eXtensible Markup Language.

XML schema – The definition of an XML document, specifying required and optional XML

elements, their order, and parent-child relationships.

XML Schematron – A set of rules used to validate an XML document.

1 Overview

1.1 Purpose and Scope

This software interface specification (SIS) describes the format and content of the bundles containing Camera data for the Mars 2020 Project.

This document specifically describes the PDS data archive bundle and collections in which data products, documentation, and supporting material are stored. This document is intended for the scientists who will analyze the data, including those associated with the project and those in the general planetary science community. It describes how the data is organized within the PDS bundle. It does *not* describe the products themselves. For details about the products, see the Mars 2020 Software Interface Specification (SIS): Camera Instrument Experiment Data Record (EDR) and Reduced Data Record (RDR) Data Products (called "Camera SIS, see Applicable Document 5).

1.2 SIS Contents

This SIS discusses standards used in generating the data products and software that may be used to access the products. The data structure and organization are described in sufficient detail to enable a user to locate desired data products and their associated files within the archive bundle.

Appendices include a list of cognizant persons involved in generating the archive.

1.3 Applicable Documents

1. Planetary Data System Standards Reference, version 1.16.0.0, April 21, 2021.

2. Planetary Data System (PDS) 4 Data Dictionary Document, Abridged, version 1.16.0.0, April, 2021.

3. Planetary Data System (PDS) PDS4 Information Model Specification, version 1.16.0.0, April, 2021.

4. Data Provider's Handbook, Archiving Guide to the PDS4 Data Standards, version 1.16.0.0, April 21, 2021.

5. Mars 2020 Software Interface Specification: Camera Instrument Experiment Data Record (EDR) and Reduced Data Record (RDR) Data Products, version 1.0, Sep 1, 2020. https://pds-geosciences.wustl.edu/m2020/urn-nasa-pds-mars2020_mission/document_camera/Mars2020_Camera_SIS.pdf

The PDS4 Documents 1 through 4 are subject to revision. Specifically, these are the versions current as of release 2. Subsequent release may use newer versions of the IM; this is described in the label (information_model_version attribute). The most recent versions may be found at http://pds.nasa.gov/pds4. Document 5 is included as part of this archive; the version in the archive may be more up to date and should be used in preference to the version listed above.

1.4 **Audience**

This document serves as an Archive Bundle SIS, describing the structure and content of the archive in which the data products, documentation, and supporting material are stored. It does not describe the format and content of Mars 2020 camera data products in detail – that information is addressed in the Camera SIS (Applicable Document 5). This SIS is intended to be used both by the instrument teams in generating the archive, and by data users wishing to understand the format and content of the archive. Typically these individuals would include scientists, data analysts, and software engineers.

2 Mars 2020 Camera Data Products

2.1 Data Product Overview

The Mars 2020 camera raw and derived products are described in detail in the Mars 2020 Camera SIS (Applicable Document 5).

2.2 Data Processing Levels

Data processing levels mentioned in this SIS refer to PDS4 processing levels. Table 2-1 provides a description of these levels along with the equivalent designations used historically in other systems, particularly when describing data products for cameras on landed/rover missions. Note that the data processing levels are not used as an organizational principle for the camera bundles and collections, despite some name similarities (see Section 3.1.1).

PDS4 processing level	PDS4 processing level description	NASA Level (used in PDS3)
Telemetry	Telemetry data with instrument data embedded. PDS does not archive telemetry data.	
Raw	Original data from an instrument. If compression, reformatting, packetization, or other translation has been applied to facilitate data transmission or storage, those processes are reversed so that the archived data are in a PDS approved archive format. Often called EDRs (Experimental Data Records).	0
Partially Processed	Data that have been processed beyond the raw stage but which have not yet reached calibrated status. These and more highly processed products are often called RDRs (Reduced Data Records), but some EDRs are also in this category.	1A
Calibrated	Data converted to physical units, which makes values independent of the instrument. Often called RDRs.	1B
Derived	Results that have been distilled from one or more calibrated data products (for example, maps, gravity or magnetic fields, or ring particle size distributions). Supplementary data, such as calibration tables or tables of viewing geometry, used to interpret observational data should also be classified as 'derived' data if not easily matched to one of the other three categories. Often called RDRs.	2+

Table 2-1: Data processing level definitions

3 Camera Archive Organization, Identifiers and Naming Conventions

This section describes the basic organization of the Mars 2020 camera raw, partially processed, calibrated and derived data archived under the PDS4 Information Model (IM) (Applicable Documents 1 through 4), including the naming conventions used for the bundle, collection, and product unique identifiers.

3.1 The Mars 2020 Camera Bundles

The highest level of organization for a PDS4 archive is the bundle. A bundle is a set of one or more related collections which may be of different types. A collection is a set of one or more related basic products which are all of the same type. Bundles and collections are logical structures, not necessarily tied to any physical directory structure or organization. In the following paragraphs the term "instrument" refers to the science or engineering instruments on the rover, which may produce images, other types of scientific data, or both, while "camera" refers to the portion of an instrument that produces images.

The Mars 2020 camera data is divided into many bundles, as listed in Table 3-1. There are a lot of bundles; Table 3-2 should help make sense of them. There are three different bundle configurations: standalone camera bundles, secondary instrument bundles containing image operations camera data, and primary instrument bundles. The relationships between the bundles is illustrated in Figure 1.

The mars2020_imgops bundle contains the camera data for the SHERLOC (ACI and Watson), PIXL (MCC), and SuperCam (RMI) cameras, as processed by the Mars 2020 image operations team (IDS – Instrument Data System). The collections in this bundle are included *by reference* as secondary collections in the respective instrument bundles (mars2020_sherloc, mars2020_pixl, mars2020_supercam) using the PDS "secondary collection" mechanism. Thus the data logically appears both the camera "imgops" bundle and in the instrument-specific bundle, but it is physically only stored in the "imgops" bundle. The imgops bundle may also include camera data processed by the SHERLOC team; camera data processed by the PIXL or SuperCam teams appears in their respective instrument bundles (only). See the Bundle SIS's for those bundles for details.

The MEDA Skycam data appears in the mars2020_meda instrument data bundle. It is fully part of the MEDA bundle (described by the MEDA Bundle SIS) but is also described here due to its similarity to all the other cameras' data.

All of the bundles described here except MEDA are maintained at the PDS Cartography and Imaging Sciences node. MEDA is maintained at the Atmospheres node. The instrument bundles containing the non-camera data for SHERLOC, PIXL, and SuperCam are maintained at the Geosciences node.

3.1.1 Release 2+ Bundle Structure Revision

In the first release of Mars2020 data to PDS, there was one bundle for all of the Mastcam-Z data, and one bundle for all of the Navcam, Hazcam, EDLCAM, and Cachecam data. Because of the sheer data volume of those bundles, starting in release 2 they were broken up into several different bundles. Table 3-2 shows the revised (release 2+) structure. The bundles appearing only in release 1 are shown in Table 3-3.

It is important to note that all relevant products (those not superceded by new versions or renamed in release 2) from release 1 have been included as secondary products products in the release 2 bundles. Thus, the revised bundles (Table 3-1 and Table 3-2) contain all of the mission data.

The new bundles are divided into sets by instrument, and team that produced them. For the Enginering cameras, that means the bundle has been split into sets for each of Navcam, Hazcam, Cachecam, and EDLCAM (all are produced by JPL-IDS). For Mastcam-Z, the bundles have been split into two sets: data generated by the JPL-IDS team (`_ops`) and data generated by the Mastcam-Z science team (`_sci`).

The name "ops" in the bundle names is shorthand for data generated by the "Operations Team", namely JPL-IDS. The existence of a "sci" bundle in some cases in no way implies that the "ops" bundles are not valid for scientific purposes. They are simply produced by different teams, with different algorithms, as described in the Camera SIS. It is up to the user which data they want to use, based on their needs.

3.1.2 Bundle Sets

As a result of this bundle structure revision, there are a number of actual bundles for each of the engineering camera and Mastcam-Z bundle "sets". These sets are indicated by a suffix to the base name, as shown in Table 3-2. The data contents within each suffix are consistent across all these bundle sets, and are described below. Note that "bundle sets" is not a PDS term; it is used informally here to group related bundles for user convenience.

It is very important to note, the names "raw" and "calibrated" here should be interpreted in the *generic* sense of the words – **not** as PDS Processing Levels. As a matter of fact, they contain data from several different PDS Processing Levels.

- No suffix: all camera data for the instrument(s) is in the single bundle
- `_raw`: Contains all data in the EDR class (raw, decompanded, and de-Bayered)
- `_calibrated`: Contains data that can be processed from a single frame, with no partner images of any kind. Generally this is radiometric and color corrected data, as well as the TDR and FDR classes of data.
- `_stereo`: Contains data that is the result of stereo processing (broadly defined, including depth-from-focus or other such techniques).
- `_mesh`: Contains terrain meshes (3D models) derived from stereo data.
- `_mosaic`: Contains mosaics in various projections.

- `_video`: Contains MPEG videos from the EDLCAM instrument
- `_audio`: Contains audio files from the EDLCAM instrument

The specific contents of the Raw, Calibrated, and Stereo are defined below, using the categories defined in Table 17-2 of the Camera SIS (Applicable Document 5). Cross-referencing the tables will show which bundle any given 3-letter type code is in.

- EDR: `raw`
- FDR: `calibrated`
- Radiometric: `calibrated`
- Color: `calibrated`
- Correlation: `stereo`
- XYZ: `stereo` (except MXY)
- Range: `stereo`
- Surface Normal: `stereo`
- Slope: `stereo`
- Reachability: `stereo`
- Roughness: `stereo`
- Goodness: `stereo`
- Miscellaneous: varies
    - MSK: `calibrated`
    - TEN/TER: `stereo`
    - IEF/IEP: `stereo`
    - IDM/IDX/ICM: `calibrated`

There is one exception: the MXY type (in the XYZ category) is in calibrated, because it can be generated without stereo data.

Sorted the other direction for convenience:

- `raw`: EDR
- `calibrated`: FDR, Radiometric, Color, MXY, MSK, IDM/IDX/ICM
- `stereo`: Correlation, XYZ, Range, Surface Normal, Slope, Reachability, Roughness, Goodness, TEN/TER, IEF/IEP

Note that the "`mosaic`" bundles contain all mosaics, regardless of whether they're made out of images from the `raw`, `calibrated`, or `stereo` bundles. Similarly, the "`mesh`" bundle contains all meshes.

Camera Bundle Name / Bundle Set Base Name (PDS Node)	Non-Camera Instrument Bundle Name (PDS Node)	Image and Browse Collection Names	Image Data Description
mars2020_navcam_ops_* (IMG)	n/a	data browse	Navcam data produced by the JPL-IDS team
mars2020_hazcam_ops_* (IMG)	n/a	data browse	Hazcam data produced by the JPL-IDS team
mars2020_cachecam_ops_* (IMG)	n/a	data browse	Cachecam data produced by the JPL-IDS team
mars2020_edlcam_ops_* (IMG)	n/a	data browse	EDLCAM data produced by the JPL-IDS team
mars2020_mastcamz_ops_* (IMG)	n/a	data browse	Mastcam-Z data produced by the JPL-IDS team
mars2020_mastcamz_sci_* (IMG)	n/a	data browse	Mastcam-Z data produced by the ASU instrument team
mars2020_helicam (IMG)	n/a	data browse	Helicopter Camera data (Heli Nav, Heli RTE)
mars2020_imgops (IMG)	mars2020_sherloc (GEO)	data_watson_imgops browse	SHERLOC Watson data, from both JPL-IDS and instrument team
mars2020_imgops (IMG)	mars2020_sherloc (GEO)	data_aci_imgops browse	SHERLOC ACI data, from both JPL-IDS and instrument team
mars2020_imgops (IMG)	mars2020_pixl (GEO)	data_mcc_imgops browse	PIXL MCC data, from JPL-IDS
mars2020_imgops (IMG)	mars2020_supercam (GEO)	data_rmi_imgops browse	SuperCam RMI data, from JPL-IDS
mars2020_meda (ATM)	mars2020_meda (ATM)	data_skycam browse_skycam	MEDA SkyCam data, from JPL-IDS. Bundle also contains the rest of the MEDA data.

*Table 3-1: Mars 2020 Camera Bundles and Bundle Sets (release 2+). See Table 3-2 for the meaning of the *'s for the bundle sets.*

Bundle Set Root Name	No suffix	_raw	_calibrated	_stereo	_mesh	_mosaic	_video	_audio
mars2020_navcam_ops	-	X	X	X	X	X	-	-
mars2020_hazcam_ops	-	X	X	X	X	X	-	-
mars2020_cachecam_ops	-	X	X	-	-	-	-	-
mars2020_edlcam_ops	-	X	X	-	-	-	X	X
mars2020_mastcamz_ops	-	X	X	X	X	X	-	-
mars2020_mastcamz_sci	-	-	X	-	-	-	-	-
mars2020_helicam	X	-	-	-	-	-	-	-
mars2020_imgops	X	-	-	-	-	-	-	-
mars2020_meda	X	-	-	-	-	-	-	-

Table 3-2: Mars 2020 Camera Set Bundle Suffix Names (release 2+)

Camera Bundle Name (PDS Node)	Non-Camera Instrument Bundle Name (PDS Node)	Image and Browse Collection Names	Image Data Description
mars2020_ecam (IMG)	n/a	data browse	Engineering Camera data (Hazcams, Navcams, EDL cameras, LVS camera, Cachecam)
mars2020_mastcamz (IMG)	n/a	data_ids browse_ids	Mastcam-Z data produced by the IDS team
mars2020_mastcamz (IMG)	n/a	data_asu browse_asu	Mastcam-Z data produced by the ASU instrument team

Table 3-3: Mars 2020 Camera Bundles for release 1 that are deprecated in release 2

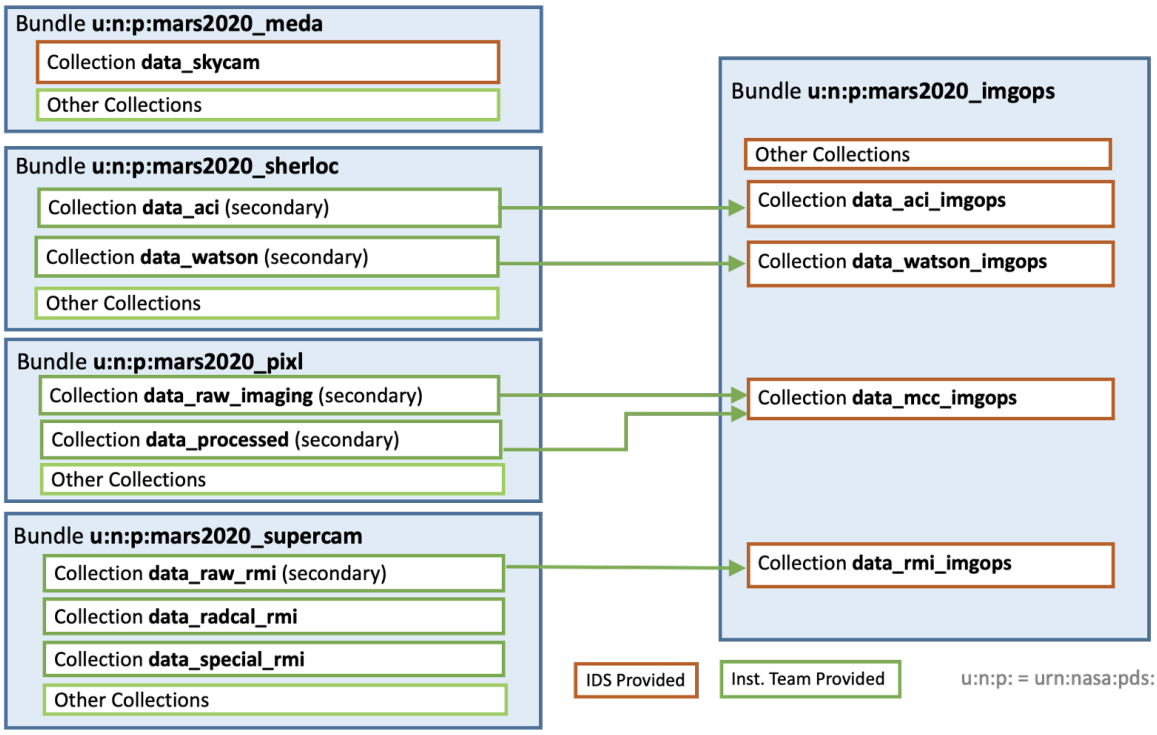

Figure 1. Camera and Instrument Bundle and Collection Relationships

3.2 Mars 2020 Mission bundle

There is one additional bundle that is important: the mars2020_mission bundle. This bundle is the primary repository for documents, calibration data, and other items that are relevant across multiple instrument or camera bundles.

This means that most of the items in the Camera Bundles relating to documentation, calibration, and Velocity templates ("miscellaneous") actually reside in the Mission bundle, with the various Camera bundles including the collections as secondary collections. This means for example the Camera bundles each have a pointer to the Camera SIS that actually resides in the Mission document. In this way, there is only one copy of such items, while still being a logical member of each bundle.

Note that bundles may have their own specific "document" collection with documents applicable only to that bundle. The "mars2020_mastcamz_sci_calibrated" bundle is an example of this.

While the Mission bundle has its own SIS, an outline is shown in Table 3-4 for the parts that are relevant to the camera bundles.

Mission Bundle Collection Name	Camera Bundle Collection Name	Contents
document	n/a	General documentation for the mission, not specific to cameras

Mission Bundle Collection Name	Camera Bundle Collection Name	Contents
document_camera	document_camera	Camera-related documents: this document, Camera Data Product SIS and related files
calibration_camera	calibration_camera	Camera calibration files
miscellaneous	miscellaneous	Velocity templates used to generate the PDS4 labels

Table 3-4: Mission Bundle Contents

3.2.1 Velocity Templates in the Miscellaneous Collection

The Miscellaneous collection (in the Mission bundle, or as a secondary collection in the other bundles) contains the Velocity templates used to make the PDS4 labels out of the VICAR or ODL labels associated with the data products.

Velocity is a Java-based template engine developed by the Apache project (https://velocity.apache.org). It is used by the PDS tool "mi-label" (https://nasa-pds.github.io/mi-label/) to create a powerful language for building PDS4 XML labels. See the Apache documentation for details, but in summary, the template looks a lot like the XML it is supposed to produce, where "$" is for variable substitution, "#" is for control structures and macro (subroutine) calls, and "##" is for comments.

It is important to note that these templates are not part of the science data. They are supplied in case they are helpful for users to understand the structure of the labels, particularly how VICAR/ODL and PDS4 labels relate to each other. They are not needed in any way in order to make use of the data.

3.3 **Collections in the Camera Bundles**

Collections consist of basic products that are related. The Mars 2020 Camera Bundles each include the data collections listed in Table 3-5 (below). These collections include both the official scientific products, and supplemental products such as FITS files (SuperCam) or the original JPEG stream.

The name of the data collection varies per bundle; see Table 3-1.

The data collection for images contains any or all of the PDS processing levels: Raw, Partially Processed, Calibrated, and Derived. There is no association between the processing level and the bundle or collection names (despite some similarity in names). The Camera SIS lists what the processing levels are for each product type.

Collection Logical Name	Collection Type	Description
See Table 3-1, column 3	Browse	"Quick look", possibly compressed versions of the science data products. The browse versions of data products are not science quality.
calibration_camera	Calibration	Secondary collection of calibration reports and files needed for calibration of camera data. The primary collection is: u:n:p:mars2020_mission:calibration_camera
See Table 3-1, column 3	Data	The actual scientific data products for the camera. See the Data Organization section for a description of how the data will be organized into sub-directories.
document_camera	Document	Secondary collection of documents which are useful for understanding and using the data. This collection includes the Camera SIS (Applicable Document 5), and any additional documentation the Mars 2020 team wishes to include. Documentation files are PDF-A, plain text or HTML format to be PDS-compliant. The primary collection is: u:n:p:mars2020_mission:document_camera
document	Document	Primary collection containing documents specific to an instrument. Only exists in some bundles.
miscellaneous	Miscellaneous	Secondary collection of Apache Velocity templates used to generate the PDS4 labels from the VICAR/ODL header labels. The primary collection is: u:n:p:mars2020_mission:miscellaneous

Table 3-5: Collections in the Mars 2020 Camera Bundles

3.4 **Data Organization**

The data directory contains all files from the data collection. These are organized by sol, category, and instrument. The data directory organization directly mirrors the operational directory structure used by Mars 2020, which facilitates use of operations tools on the PDS archive, as well as the creation and validation of the archive. See Appendix B for the full directory structure of the bundles, outside of the data directory. All of the camera bundles use the same organization for their data collections. Particularly in the case of the engineering camera and Mastcam-Z bundle sets, this can lead to sparse-looking directories with only one subdirectory. This is on purpose: the intent is that users could merge all the data and browse bundles across all Mars 2020 bundles into one physical directory structure and not have any name collisions – the result would be a unified directory structure containing all the data.

The data collection directory structure is:

```
.. <data_collection_name>/
.... sol/
...... <sol_number>/
........ <team>/
.......... <category>/
............ <instrument>/
```

Where:

"data_collection_name" is the name of the collection, from Table 3-1.

"sol_number" is the Mars day since landing (landing day is sol 0). The sol_number is a 5-digit, zero-filled, decimal number (e.g. 00001).

"team" is the name of the team generating the products. The name is "ids" for products generated by the IDS team at JPL. The name "soas" is used for products generated by the science teams (Mastcam-Z and SHERLOC).

"category" is one of: edr, fdr, rdr, tdr, where edr contains all data products in the EDR category (raw and some partially processed); fdr contains the "fundamental" data record, which is a consistently formatted product regardless of how the image was downlinked; rdr contains most Partially Processed, all Calibrated, and Derived products, and tdr is analogous to fdr for engineering camera tile data (see the camera SIS). For the bundle sets, those bundles named *_raw contain only edr; those named *_calibrated contain fdr, tdr, and rdr, and the rest contain only rdr.

"instrument" is the instrument, which is one of: cachecam, edl, fcam, heli, meda, ncam, pixl, rcam, shrlc, scam, zcam, mosaic, mesh. Most of those are instrument types. "mosaic" contains all mosaics for any instrument, while "mesh" contains all terrain meshes for any instrument. Note that mosaics or meshes can be made up of data from multiple instruments; the data will go in the appropriate bundle for one of the instruments, but which is up to the discretion of the data provider, in consulation with the instrument teams.

3.5 Products in the Mars 2020 Bundles

A PDS product consists of one or more digital objects and an accompanying PDS4 label file, which provides identification and description information for labeled objects. Documents are also considered products, and have PDS4 labels just as data products do. In addition, collections and even the bundles are considered products and have PDS4 labels.

3.5.1 Logical Identifiers

Every product in PDS is assigned an identifier, which allows it to be uniquely identified across the system. This identifier is referred to as a Logical Identifier (LID). A LIDVID (Versioned

Logical Identifier) includes product version information, and allows different versions of a specific product to be referenced uniquely. A product's LID and VID (Version Identifier) are defined as separate attributes in the product label. LIDs and VIDs are assigned by PDS and are formed according to the conventions described in the LID Formation and VID Formation sections below. The uniqueness of a product's LIDVID may be verified using the PDS Registry and Harvest tools.

3.5.2 LID Formation

LIDs take the form of a Uniform Resource Name (URN). LIDs are restricted to ASCII lower case letters, digits, dash, underscore, and period. Colons are also used, but only to separate prescribed components of the LID. Within one of these prescribed components dash, underscore, or period are used as separators. LIDs are limited in length to 255 characters.

Mars 2020 Camera LIDs are formed according to the following conventions:

- Bundle LIDs are formed by appending a bundle specific ID to the base ID:

urn:nasa:pds:<bundle ID>

Example: urn:nasa:pds:mars2020_navcam_ops_stereo

The <bundle ID> must be unique across all PDS data archive bundles. See Table 3-1, Table 3-2, and Figure 1.

- Collection LIDs are formed by appending a collection specific ID to the collection's parent bundle LID:

urn:nasa:pds: <bundle ID>:<collection ID>

Example: urn:nasa:pds:mars2020_navcam_ops_stereo:data

Because the collection LID is based on the bundle LID, the only syntactic condition is that the <collection ID> must be unique across the bundle. Thus collection LIDs are unique across PDS. Collection IDs correspond to the collection type (e.g. "browse", "data", "document", etc.). See Table 3-5 for examples of Collection IDs.

- Basic product LIDs are formed by appending a product specific ID to the product's parent collection LID:

urn:nasa:pds: <bundle ID>:<collection ID>:<product ID>

Because the product LID is based on the collection LID, the only additional syntactic condition is that the <product ID> must be unique across the collection. However, for the Mars 2020 camera data, product IDs are unique across the entire set of Camera bundles.

For the Mars 2020 data and browse collections, the <product ID> is generated as follows:

- Convert the product filename to lowercase.
- Remove the version number if present (last two characters of filename, before extension).
- Remove the .VIC or .IMG extension for images only.

Note that non-image files and image files other than .VIC or .IMG (e.g. .PNG) retain their extension to preserve uniqueness across multiple file types using the same base name (e.g. mosaic ancillary files or mesh OBJ files).

For example:

Filename:

`NLG_0022T0668847813_870ECM_N0010000NCAM00160_01_2I3J01.IMG`

LID :

`urn:nasa:pds:mars2020_navcam_ops_raw:data:`
`nlg_0022t0668847813_870ecm_n0010000ncam00160_01_2i3j`

Filename:

`ZR0_0200T0648036352_101ECM_N0010000XXXX00000_028100J01.obj`

LID:

`urn:nasa:pds:mars2020_mastcamz_ops_raw:data:`
`zr0_0200t0648036352_101ecm_n0010000xxxx00000_028100j.obj`

For the `calibration_camera` collection, the same rules apply, except there is no version number to remove.

For the remaining collections, the <product ID> is set to be the same as the data file name, with the extension, except for IMG or VIC files, where the extension is stripped.

3.5.3 VID Formation and Product Versioning

Product Version IDs consist of major and minor components separated by a "." (M.n). Both components of the VID are integer values. The major component is set based on the operations pipeline version ID assigned to the product. This value is extracted from the last two characters of the ops product filename. For Mars 2020 cameras, the minor version is always "0". Thus the complete VID for all Mars 2020 camera products is "m.0".

It is important to note that the version numbers used for archive are completely consistent with those used in operations. This facilitates scientific traceability – if someone publishes a paper using a filename from ops (a not infrequent occurrence), that file cannot be confused with a different product that exists in the archive. The ops file may or may not be available in the archive, but it will be clear to users if the file is NOT in the archive – they will not confuse it

with a different file in the archive using the same version. The intent is not to make all operations products available to the archive, but rather to ensure that there is no confusion regarding any given product. If it has not been archived, the user can definitely *know* that is the case. Often, the specific version does not matter, so in many cases finding the version that is available in the archive is sufficient. But at least the user is aware of this, in cases where it does matter.

This naturally leads to skipped version numbers in the archive. The archive might have versions 2 and 5 only, with no version 1, 3, or 4. This is somewhat at odds with historical PDS usage where a separate versioning namespace is used for PDS than from ops, but it is the only way to ensure that users can properly trace scientific results to their source products.

3.5.4 File Naming Convention

The file naming convention for Mars 2020 camera products is described in the File Naming section of the Camera SIS (Applicable Document 5).

3.5.5 PDS4 Labels

Each Mars 2020 camera data product has an accompanying PDS4 label. PDS4 labels are ASCII text files written in the eXtensible Markup Language (XML). Product labels are detached from the files they describe. There is one label for every product. A product, however, may consist of one or more data objects. The data objects of a given product may all reside in a single file, or they may be stored in multiple separate files, in which case the PDS4 label points to all the files. An example of this in Mars 2020 is mesh products, which consist of a tuple of (.obj, .mtl, .png) files, with a single (.xml) label file. Another example is JPEG EDRs, which consist of the primary data file (ECM type) and a copy of the original JPEG stream as sent by the spacecraft (EJP type) as a supplemental product. A PDS4 label file usually has the same name as the data product it describes, but always with the extension ".xml". In the case of images (.VIC or .IMG) and documentation/bundle files, ".xml" will replace the filename extension (e.g. PRODUCT_A.IMG will have the label file PRODUCT_A.xml). For most other files, the ".xml" is appended (e.g. PRODUCT_A.obj has the label PRODUCT_A.obj.xml). The `calibration_camera` collection has special rules, see Appendix C in the Camera SIS (Applicable Document 5).

The Mars 2020 image data files produced by the JPL-IDS team are in VICAR format with attached labels in both VICAR and ODL format (see Section 3.2 of the Camera SIS (Applicable Document 5)). Those produced by the instrument teams may or may not have the VICAR label, but will have an ODL label. The metadata found in these VICAR labels (or ODL if not available) are used as the inputs to generate both the ODL and the PDS4 labels, which are detached XML files. The images are thus simultaneously valid as both VICAR (data file only), ODL (data file only; PDS3 format without specifically being approved as PDS3), and PDS4 (data file plus label) products. Because the VICAR label is used to generate the ODL label, which is then used to generate the PDS4 label, the semantic contents of all three labels are identical (possible exceptions are called out in the Release Notes).

Documents are also considered products, and have accompanying PDS4 labels.

For the Mars 2020 mission, the structure and content of PDS labels will conform to the PDS core schema and Schematron based upon the PDS Information Model (Applicable Document 3). In brief, the Schema is the XML model that PDS4 labels must follow, and the Schematron is a set of validation rules that are applied to PDS4 labels. In addition to the PDS core schema documents, there are discipline- and mission-specific XML schema and Schematron documents, which provide additional governance over the products in this archive. The PDS Validate Tool (https://pds.nasa.gov/pds4/software/validate) should be used to validate the structure and content of the product labels. In brief, the Schema is the XML model that PDS4 labels must follow, and the Schematron is a set of validation rules that are applied to PDS4 labels.

A list of the XML Schema and Schematron documents associated with this archive are provided in Table 3-6. Also shown is the namespace used in the label when referencing that dictionary.

Namespace	XML Document	Steward	Product LID
(n/a)	PDS Core Schema/Schematron	PDS	urn:nasa:pds:system_bundle:xml_schema:pds-xml_schema
img:	Imaging Dictionary Schema/Schematron	Imaging Node	urn:nasa:pds:system_bundle:xml_schema:img-xml_schema
geom:	Geometry Dictionary Schema/Schematron	Geosciences Node	urn:nasa:pds:system_bundle:xml_schema:geom-xml_schema
cart:	Cartography Dictionary Schema/ Schematron	Imaging Node	urn:nasa:pds:system_bundle:xml_schema:cart-xml_schema
proc:	Processing Information Dictionary Schema/ Schematron	Imaging Node	urn:nasa:pds:system_bundle:xml_schema:proc-xml_schema
disp:	Display Information Dictionary Schema/Schematron	Imaging Node	urn:nasa:pds:system_bundle:xml_schema:disp-xml_schema
msn:	Mission Information Dictionary Schema/Schematron	Imaging Node	urn:nasa:pds:system_bundle:xml_schema:msn-xml_schema
msn_surface:	Surface Mission Information Dictionary Schema/Schematron	Imaging Node	urn:nasa:pds:system_bundle:xml_schema:msn_surface-xml_schema
img_surface:	Surface Imaging Dictionary Schema/ Schematron	Imaging Node	urn:nasa:pds:system_bundle:xml_schema:img_surface-xml_schema
msss_cam_mh:	MSSS Camera Mini-Header Dictionary Schema/Schematron	Imaging Node	urn:nasa:pds:system_bundle:xml_schema:msss_cam_mh-xml_schema
mars2020:	Mars 2020 Misson-Specific Dictionary Schema/Schematron	Geosciences Node	urn:nasa:pds:system_bundle:xml_schema:mars2020-xml_schema

Table 3-6: Mars 2020 Data Dictionaries

4 Mars 2020 Archive Bundle Product Formats

Data that comprise the Mars 2020 raw data archive are formatted in accordance with PDS specifications (see Applicable Documents 1 through 4). This section provides details on the formats used for each of the products included in the archive.

4.1 Science Data Product Formats

The telemetry, raw, partially processed, calibrated and derived data products are described in the Mars 2020 Camera SIS (Applicable Document 5).

4.2 Document Product Formats

Documents in this archive are provided as PDF/A (www.pdfa.org/download/pdfa-in-a-nutshell) or as plain ASCII text if no special formatting is required. Figures that accompany documents are provided as TIFF, GIF, JPEG, or PNG files. HTML versions of some documents are provided in addition to the PDF/A version.

4.3 Calibration Product Formats

Products in the Calibration collection include documents (PDF/A or plain ASCII text), and image files, similar in format to the science data products.

4.4 Browse Product Formats

Products in the Browse collection do not contain any scientifically useful information that is not found in the Data collection, and conform to a well-documented standard in current use (PNG). This data is not archival data, and thus may become outdated.

Appendix A Support staff and cognizant persons

Mars 2020 IDS Team			
Name	Affiliation	Phone	Email
Robert Deen	IDS Data Scientist	818-354-7492	Bob.Deen@jpl.nasa.gov
Hallie Abarca	IDS Operations Lead	818-354-1817	Hallie.E.Gengl@jpl.nasa.gov
Stirling Algermissen	IDS Development Lead	818-394-3183	Stirling.S.Algermissen@jpl.nasa.gov

PDS Imaging Node			
Name	Affiliation	Phone	Email
Myche McAuley	Imaging Node, Co-I	818-354-9478	michael.mcauley@jpl.nasa.gov
Trent Hare	Imaging Node, Mission Interface	928-556-7126	thare@usgs.gov

PDS Geociences Node			
Name	Affiliation	Phone	Email
Susie Slavney	Washington University	314-935-9295	susan.slavney@wustl.edu

PDS Atmospheres Node			
Name	Affiliation	Phone	Email
Lyle Huber	New Mexico State University	575-646-1862	lhuber@nmsu.edu

Table 4-1: Archive Support Staff

Appendix B Bundle Directory Structure

This section provides a simple outline of the bundle directory structure.

```
<bundle_name>
├── <browse_collection_name>
│    └── sol
│        └── <sol_number>
│           └── <team>
│              └── edr
│                 └── <instrument>
│              └── fdr
│                 └── <instrument>
│              └── rdr
│                 └── mosaic
│                    └── <instrument>
├── <data_collection_name>
│    └── sol
│        └── <sol_number>
│           └── <team>  ids | soas
│               └── <category>  edr | fdr | tdr | rdr
│                  └── <instrument>
│                  └── mosaic
│                  └── mesh
├── calibration_camera  (secondary collection)
├── document_camera  (secondary collection)
├── document  (primary collection, some bundles only)
├── miscellaneous  (secondary collection)
```

For example:

```
mars2020_imgops
|readme.txt
├── browse
|    └── sol
|        └── 00620
|            └── ids
|                └── edr
|                    └── shrlc
|                    └── pixl
|                    └── scam
|                └── fdr
|                    └── shrlc
|                    └── pixl
|                    └── scam
|                └── rdr
|                    └── mosaic
|                    └── shrlc
|                    └── pixl
|                    └── scam
├── data_rmi_imgops
|    └── sol
|        └── 00620
|            └── ids
|               └── edr
|                   └── scam
|               └── rdr
|                   └── scam
|                   └── mesh
|                   └── mosaic
├── data_watson_imgops
|    └── sol
|        └── 00620
|            └── ids
|               └── edr
|                   └── shrlc
|               └── rdr
|                   └── shrlc
|                   └── mesh
|                   └── mosaic
├── data_aci_imgops
```

```
  │    └── sol
  │        └── 00620
  │            └── ids
  │                └── edr
  │                    └── shrlc
  │                └── rdr
  │                    └── shrlc
  │                    └── mesh
  │                    └── mosaic
├── data_mcc_imgops
  │    └── sol
  │        └── 00620
  │            └── ids
  │                └── edr
  │                    └── pixl
  │                └── rdr
  │                    └── pixl
  │                    └── mesh
  │                    └── mosaic
├── document_camera  (secondary collection)
├── calibration_camera  (secondary collection)
├── miscellaneous  (secondary collection)
```