

Mars 2020 Project

Software Interface Specification (SIS)

Entry, Descent and Landing Camera Raw Archive Bundle Structure

Version 1.1

Release Date: November 1, 2021

JPL Doc ID: TBS

Prepared by: Amy Culver, Nicholas Ruoff

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.

(c) 2020 California Institute of Technology. Government sponsorship acknowledged.

TABLE OF CONTENTS

1	Overview	8
1.1	Purpose and Scope	8
1.2	SIS Contents	8
1.3	Applicable Documents	8
1.4	Audience	9
2	Mars 2020 EDL Camera and Microphone Data Products.....	10
2.1	Data Product Overview	10
2.2	Data Processing Levels.....	10
3	Archive Organization, Identifiers and Naming Conventions	11
3.1	The Mars 2020 EDL Camera & Microphone Raw Data Bundle	11
3.2	Relationship to other Mars 2020 bundles.....	11
3.2.1	ECAM Camera bundle.....	11
3.2.2	Mission bundle.....	11
3.3	Collections in the Camera Bundles.....	12
3.4	Products in the Mars 2020 Bundles	12
3.4.1	Logical Identifiers.....	13
3.4.2	LID Formation	13
3.4.3	VID Formation.....	13
3.4.4	File Naming Convention	14
3.4.5	PDS4 Labels	15
4	Mars 2020 Archive Bundle Product Formats	17
4.1	Science Data Product Formats	17
4.2	Document Product Formats	17
Appendix A	Support staff and cognizant persons	18

DOCUMENT CHANGE LOG

Version	Change	Date	Affected portion
1.0	Initial release	2021-07-21	All
1.1	Remove list of concatenated audio files delivered in Release 1	2021-10-21	3.4.4.2 Concatenated Audio Files

TBD ITEMS

Item	Section(s)	Page(s)

ACRONYMS AND ABBREVIATIONS

Acronym	Meaning
ACI	Autofocusing Context Imager (SHERLOC camera)
ASCII	American Standard Code for Information Interchange
DSU	Data Storage Unit
EDL	Entry, Descent and Landing
EDR	Experiment Data Record
FDR	Fundamental Data Record
HTML	Hypertext Markup Language
IDS	Instrument Data System (at JPL)
IM	Information Model
JPL	Jet Propulsion Laboratory
LID	Logical Identifier
LIDVID	Versioned Logical Identifier
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	Planetary Instrument for X-ray Lithochemistry (Mars 2020 instrument)
RDR	Reduced Data Record
RTE	Return to Earth (Helicopter camera)
SOAS	Science Operations Analysis Subsystem

Mars 2020 Camera Archive Bundle Software Interface Specification

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
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 – often used in plural. The term may be capitalized when referring 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.

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.

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 Information Model, 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 bundle containing raw data for Entry, Descent and Landing (EDL) cameras and microphone 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 1.14.0, May 22, 2020.
2. Planetary Data System (PDS) 4 Data Dictionary Document, Abridged, version 1.14.0.0, June, 2020.
3. Planetary Data System (PDS) PDS4 Information Model Specification, version 1.14.0.0, June, 2020.
4. Data Provider’s Handbook, Archiving Guide to the PDS4 Data Standards, version 1.14.0, May 19, 2020.
5. Mars 2020 Software Interface Specification: Camera Instrument Experiment Data Record (EDR) and Reduced Data Record (RDR) Data Products, latest version.

The PDS4 Documents 1 through 4 are subject to revision. The most recent versions may be found at <http://pds.nasa.gov/pds4>. Document 5 is included as part of this archive.

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 EDL camera and microphone 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 EDL Camera and Microphone Data Products

2.1 Data Product Overview

The Mars 2020 EDL camera and microphone 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.

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 Archive Organization, Identifiers and Naming Conventions

This section describes the basic organization of the Mars 2020 EDL camera and microphone raw, 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 EDL Camera & Microphone Raw Data Bundle

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.

This `m2020_edlcam_raw` bundle contains the EDL Video & Microphone raw data.

3.2 Relationship to other Mars 2020 bundles

3.2.1 ECAM Camera bundle

The `mars2020_ecam` bundle is the nominal repository for raw and processed versions of the EDL camera and microphone data generated by the Mars 2020 project Instrument Data System (IDS) team. It is expected that after Release 1, the IDS team will deliver all raw and derived products for the data returned to earth from the EDL cameras and microphone for the relevant delivery period. This `m2020_edlcam_raw` bundle simply provides a mechanism to release the raw audio and video files acquired at landing for Release 1.

3.2.2 Mission bundle

The `mars2020_mission` bundle is the primary repository for documents, calibration data, and other items that are relevant across multiple instrument or camera bundles. Most of the items in the `m2020_edlcam_raw` relating to documentation actually reside in the Mission bundle as secondary members of the `m2020_edlcam_raw` bundle.

For example, the Camera SIS actually resides in the `m2020_mission` bundle document collection, although it is referenced in the document collection of this bundle. In this way, there is only one copy of such items, while still being a logical member of each bundle. While the Mission bundle has its own SIS, an outline is shown in [Table 3-1](#) for the parts that are relevant to the `m2020_edlcam_raw` bundle.

Mission Bundle Collection Name	Camera Bundle Collection Name	Contents
document	document	General documentation for the mission, not specific to cameras
document_camera	document	Camera-related documents: this document, Camera Data Product SIS and related files

Table 3-1: Mission Bundle Contents

3.3 Collections in the Camera Bundles

Collections consist of basic products that are related. The m2020_edlcam_raw bundle includes the collections listed in [Table 3-2](#) (below).

Collection Logical Name	Collection Type	Description
document	document	Primary collection containing this SIS and the bundle release notes.
data_audio	data	Primary collection containing audio files as they were transferred from the EDL Data Storage Unit (DSU) to the Mars 2020 rover for transmission to Earth.
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 includes. The primary collection is u:n:p:mars2020_mission:document_camera
document_video	document	Primary collection containing <u>video files</u> as they were transferred from the EDL Data Storage Unit (DSU) to the Mars 2020 rover for transmission to Earth. The PDS4 standards define video files as “documents” rather than data. Thus users will find the landing videos in this document collection.

Table 3-2: Collections in the Mars 2020 Bundle

3.4 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.

3.4.1 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.

3.4.2 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.

The mars2020_edlcam_raw bundle 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:m2020_edlcam_raw

The <bundle ID> is unique across all PDS data archive bundles.

- 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:m2020_edlcam_raw:data_audio

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. See [Table 3-2](#) 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.

The <product ID> is set to be the same as the data file name including the extension.

3.4.3 VID Formation

Product Version IDs consist of major and minor components separated by a "." (M.n). Both components of the VID are integer values, and for this bundle, it is always set to 1.0.

3.4.4 File Naming Convention

3.4.4.1 Video and 1-second audio files

The filename used for the EDLCAM video and 1-second audio files in the m2020_edlcam_raw bundle is illustrated in Table 3-4:

<Instrument>_<Session ID>_<Compression Quality>_<Frame ID>.<Extension>

Field	Description
Instrument	<ul style="list-style-type: none"> • ddc : descent stage downlook camera • mic : microphone • puc1 : parachute uplook camera 1 • puc2 : parachute uplook camera 2 • puc3 : parachute uplook camera 3 • rdc : rover downlook camera • ruc : rover uplook camera
_	Underscore for readability. Always set to “_”.
Session ID	<ul style="list-style-type: none"> • nnnnnnn : 0 padded session number. New session number generated when IMAGE command is sent to the DSU. <p>Note: IMAGE command is also used to command the microphone</p>
_	Underscore for readability. Always set to “_”.
Compression Quality	<ul style="list-style-type: none"> • 000 : raw (uncompressed) • 002 – 031 : jpeg quality setting (using ffmpeg) • 001 – 031 : mpeq quality setting (using ffmpeg) <p>Note: Lower number means higher quality</p>
_	Underscore for readability. Always set to “_”.
Frame ID	<ul style="list-style-type: none"> • nnnnnn : 0 padded frame identification number. Each frame in a session increments by 1.
.	Separator for filename and extension. Always set to “.”
Extension	<ul style="list-style-type: none"> • mp4 – video files (no audio) • wav – microphone recording

Table 3-4: Filename convention for video and 1-sec audio files

3.4.4.2 Concatenated Audio Files

To provide context for the 1-second audio files the recordings have been concatenated together in chronological order. These files can be identified with “_concat_” in the filename and have the four digit sol number on which the recording was acquired.

3.4.4.3 Mapping m2020_edlcam_raw files to future IDSO archive releases

In future archive releases, IDSO will generate all of the EDLCAM video, image, and audio products. At the time of this writing, the mapping of these products to future IDSO generated bundles can be done by using the basename in the m2020_edlcam_raw bundle, which are the data product basenames as stored on the DSU. The IDSO generated products will have that same basename stored the following keywords in the Ops and PDS4 labels:

- Operations Label:
 - Video: MOVIE_FILE_NAME (Identification Group)
 - Audio: RAW_FRAME_FILENAME (Identification Group)

- PDS4 Label: <msn_surface:telemetry_alternate_name>

3.4.5 PDS4 Labels

Each m2020_edlcam_raw 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 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 Mars 2020 image data files are in VICAR format with attached labels in both VICAR and ODL format (see Section 3.2 of the Camera SIS (Applicable Document 5)). The metadata found in these VICAR labels 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 (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.

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 master 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 master 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](#) 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 3-3](#). 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 Mission-Specific Dictionary Schema/Schematron	Geosciences Node	urn:nasa:pds:system_bundle:xml_schema:mars2020-xml_schema

Table 3-3: 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.

Appendix A Support staff and cognizant persons

Mars 2020 ECAM/EDLCAM Team		
Name	Affiliation	Email
Justin Maki	JPL	Justin.N.Maki@jpl.nasa.gov
Amy Culver	JPL	Amy.Culver@jpl.nasa.gov
David Kessler	JPL	David.J.Kessler@jpl.nasa.gov
Aaron Roth	JPL	Aaron.K.Roth@jpl.nasa.gov
Nick Ruoff	JPL	Nicholas.A.Ruoff@jpl.nasa.gov
Nathan Williams	JPL	Nathan.Williams@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 Geosciences Node			
Name	Affiliation	Phone	Email
Susie Slavney	Washington University	314-935-9295	susan.slavney@wustl.edu

Table 4-1: Archive Support Staff and Cognizant Persons