Mars Science Laboratory (MSL)

Chemistry and Micro-Imaging (ChemCam)

Reduced Data Archive

Software Interface Specification

Version 1.5

June 24, 2014

Susan Slavney

Planetary Data System Geosciences Node

TABLE OF CONTENTS

TABLES

DOCUMENT CHANGE LOG

Change	Date	Affected Portions
Revisions based on revised Camera SIS	10/9/12	All
Added contents of Extras directory	10/29/12	2.8
Added contents of Calib directory	11/7/12	2.6
Corrected RDR Archive contents and information	12/27/12	All
Removed section on data product sizes and delivery rates, as they are not predictable.	4/10/13	4.2
Updated Calib directory contents	5/16/13	2.6
Updated Extra directory contents	6/24/14	2.8

TBD ITEMS

Section	Description

ACRONYMS AND ABBREVIATIONS

ASCII	American Standard Code for Information Interchange
CCS	LIBS Intermediate Clean Calibrated Spectra
CD-ROM	Compact Disk - Read-Only Memory
CD-WO	Write-Once Compact Disk
ChemCam	Chemistry and Micro-Imaging
CODMAC	Committee on Data Management and Computation
CSV	Comma-Separated-Value
EDR	Experiment Data Record
HTML	Hypertext Markup Language
ICD	Interface Control Document
ISO	International Standards Organization
JPL	Jet Propulsion Laboratory
LIBS	Laser Induced Breakdown Spectrometer (ChemCam component)
MEC	LIBS Multivariate Prediction of Elemental Composition
MIPL	Multi-Mission Instrument Processing Laboratory
MOC	LIBS Multivariate Prediction of Oxide Composition
MSL	Mars Science Laboratory
NSR	LIBS Nearby Sample of Reference
NSSDC	National Space Science Data Center
ODL	Object Description Language
OPGS	Operations Product Generation Subsystem
PDF	Portable Document Format
PDS	Planetary Data System
PSG	Project Science Group
RDR	Reduced Data Record
RMI	Remote Micro-Imager (ChemCam component)
RSM	LIBS Sammon's Map
SDVT	Science Data Validation Team
SIS	Software Interface Specification
TBD	To Be Determined

UEC   LIBS Univariate Prediction of Elemental Composition

UOC   LIBS Univariate Prediction of Oxide Composition

GLOSSARY

Archive – An archive consists of one or more data sets along with all the documentation and ancillary information needed to understand and use the data. An archive is a logical construct independent of the medium on which it is stored.

Archive Volume, Archive Volume Set – A volume is a unit of media on which data products are stored; for example, one CD-ROM. An *archive volume* is a volume containing all or part of an archive; that is, data products plus documentation and ancillary files. It can also be thought of as the directory tree containing the archive. When an archive spans multiple volumes, they are called an *archive volume set*.

Catalog Information – Descriptive information about a data set (e.g. mission description, spacecraft description, instrument description), expressed in Object Description Language (ODL) which is suitable for loading into a PDS catalog.

Data Product – A labeled grouping of data resulting from a scientific observation, usually stored in one file. A product label identifies, describes, and defines the structure of the data. An example of a data product is a planetary image, a spectrum table, or a time series table.

Data Set – An accumulation of data products. A data set together with supporting documentation and ancillary files is an archive.

Standard Data Product – A data product generated in a predefined way using well-understood procedures, processed in "pipeline" fashion. Data products that are generated in a nonstandard way are sometimes called *special data products*. Usually standard products are those an instrument team is required to make, and special products are those that are made as needed or as resources permit.

1 Introduction

1.1 Purpose and Scope

This Software Interface Specification is intended to be used by those who wish to understand the format and content of the Mars Science Laboratory (MSL) Chemistry and Micro-Imaging (ChemCam) Reduced Data Record (RDR) Archive. Typically, these individuals would be software engineers, data analysts, or planetary scientists.The specifications in this document apply to all ChemCam RDR standard product archive volumes that are generated by the MSL Project.

1.2 Content Overview

The MSL ChemCam RDR archive consists of the following derived products, which are described in the MSL Camera EDR/RDR Software Interface Specification, applicable document 3. All ChemCam derived products are generated by the ChemCam Science Team.

Table 1. MSL ChemCam Derived Data Products

NASA Level	CODMAC Level	PDS Data Set ID	Product Type
1C	5	MSL-M-CHEMCAM-RMI-5-RDR-V1.0	RMI radiometrically calibrated images
1B	4	MSL-M-CHEMCAM-LIBS-4/5-RDR-V1.0	LIBS calibrated spectra (RDR) ASCII CSV table
1B	4		LIBS Intermediate Clean Calibrated Spectra (CCS) ASCII CSV table
2	5		LIBS Multivariate Prediction of Oxide Composition (MOC) ASCII CSV table
2	5		LIBS Multivariate Prediction of Elemental Composition (MEC, special product) ASCII CSV table
2	5		LIBS Univariate Prediction of Elemental Composition (UEC, special product) ASCII CSV table
2	5		LIBS Univariate Prediction of Oxide Composition (UOC, special product) ASCII CSV table
3	5		LIBS Sammon's Map (RSM, special product) ASCII CSV table

NASA and CODMAC data processing levels are described in Table 2.

Table 2. Processing Levels for Science Data Sets

NASA	CODMAC	Description
Packet data	Raw – Level 1	Telemetry data stream as received at the ground station, with science and engineering data embedded.

NASA	CODMAC	Description
Level-0	Edited – Level 2	Instrument science data (e.g., raw voltages, counts) at full resolution, time ordered, with duplicates and transmission errors removed.
Level 1-A	Calibrated - Level 3	Level 0 data that have been located in space and may have been transformed (e.g., calibrated, rearranged) in a reversible manner and packaged with needed ancillary and auxiliary data (e.g., radiances with the calibration equations applied).
Level 1-B	Resampled - Level 4	Irreversibly transformed (e.g., resampled, remapped, calibrated) values of the instrument measurements (e.g., radiances, magnetic field strength).
Level 2	Derived - Level 5	Geophysical parameters, generally derived from Level 1 data, and located in space and time commensurate with instrument location, pointing, and sampling.
Level 3	Derived - Level 5	Geophysical parameters mapped onto uniform space-time grids.

This Archive Volume Software Interface Specification (SIS) describes the format and content of the ChemCam RDR Archive. Section 2, Archive Volume Contents, describes the structure of the archive volumes and the contents of each file. Section 3, Archive Volume Format, describes the file formats used on the archive volumes. Section 4, Archive Volume Generation, describes the procedure for assembling the archive and transferring it to the Planetary Data System (PDS). Finally, Section 5, Support Staff and Cognizant Persons, lists the individuals responsible for generating and receiving the archive volumes.

1.3 Applicable Documents and Constraints

This Archive Volume SIS is intended to be consistent with the following documents:

1. Mars Exploration Program Data Management Plan, R. E. Arvidson, S. Slavney and J. Ward, Rev. 4, June 15, 2011.
2. Mars Science Laboratory Archive Generation, Validation, and Transfer Plan, J. Crisp and P. Theisinger, JPL D-35281, MSL-214-1333, May 28, 2010.
3. MSL Camera & LIBS Experiment Data Record (EDR) and Reduced Data Record (RDR) Data Products, SIS-SCI006-MSL, JPL D-38107, April 11, 2012.
4. Planetary Data System Archive Preparation Guide, Version 1.4, JPL D-31224, April 1, 2010.
5. Planetary Data System Standards Reference, February 27, 2009, Version 3.8, JPL D-7669, Part 2.
6. MSL ChemCam Science Team and PDS Geosciences Node Interface Control Document (ICD), Version 2.1, March 3, 2008.

1.4 Relationships with Other Interfaces

This Archive Volume SIS could be affected by changes to the design of the ChemCam RDR standard data products (Applicable Document 3).

2 Archive Volume Contents

This section describes the contents of the MSL ChemCam RDR Archive volumes, including the file names, file contents, file types, and organization responsible for providing the files. There is one archive volume with the PDS Volume ID MSLCCM_1001. With each delivery of ChemCam data to the PDS, the new products are added to the existing archive volume.

2.1 Root Directory Contents

Files in the Root Directory include an overview of the archive, a description of the volume for the PDS Catalog, and a list of errata or comments about the archive. The following files are contained in the Root Directory.

File Name	File Contents	File Provided By
AAREADME.TXT	Volume content and format information	Geosciences Node
ERRATA.TXT	A cumulative listing of comments and updates concerning all archive volumes published to date. More than just errata, this file is updated with release notes for each data release, including an explanation of any revisions or omissions.	ChemCam Team
VOLDESC.CAT	A description of the contents of this volume in a PDS format readable by both humans and computers	Geosciences Node

2.2 Data Directory Contents and Naming

The Data Directory is divided into subdirectories by sol (Mars day), with directory names in the form SOLnnnn (example: SOL0001). All ChemCam LIBS and RMI RDR products acquired on a given sol are in the same sol subdirectory. In the case of a product that spans more than one sol, it will be stored in the directory of the first sol covered. Each product is accompanied by its PDS label in a separate file with the same name but the extension .LBL. RMI products and labels are provided by OPGS, and LIBS products and labels are provided by the ChemCam Team.

The file naming scheme is fully explained in the MSL Camera EDR/RDR SIS, Applicable Document 3. In brief, the scheme is as follows:

<instr><config><spec><sclk><prod><geom><samp><site><drive><seqid><producer><ver>.<ext>

where

> **instr** = CL for ChemCam LIBS, CR for ChemCam RMI, CC for ChemCam generic
>
> **config** = 1-digit value from 0 to 7 (see definitions in [3])
>
> **spec** = 1-character indicator of special processing, or underscore if no special processing
>
> **sclk** = 9-character Spacecraft Clock Start Count
>
> **prod** = 3-character product type identifier (see definitions in [3])
>
> **geom** = 1-character indicator that product has had camera-induced distortion removed ("linearized") or not ("non-linearized"); see definitions in [3]
>
> **samp** = 1-character indicator of how data were sampled (see definitions in [3])

site = 3 character site location number

drive = 4 character drive location (position within site)

seqid = 4 character Category subfield and 5-digit sequence ID

producer = "M" for MIPL; "P" for ChemCam Principal Investigator; other characters A through L to be assigned to ChemCam Co-Investigators

ver = 1-character version identifier

ext = TAB for ASCII fixed-length table data, CSV for ASCII variable-length comma-separated table data, IMG or TIF for image data, DAT for binary table data, LBL for ASCII text PDS label.

2.3 Index Directory Contents

Files in the Index Directory are provided to help the user locate products in the archive. The following files are contained in the Index Directory.

File Name	File Contents	File Provided By
INDXINFO.TXT	A description of the contents of this directory	Geosciences Node
LIBSINDEX.TAB	A table listing all LIBS data products on this volume	ChemCam Team
LIBSINDEX.LBL	A PDS detached label that describes LIBSINDEX.TAB	ChemCam Team
RMIINDEX.TAB	A table listing all RMI data products on this volume	ChemCam Team
RMIINDEX.LBL	A PDS detached label that describes RMIINDEX.TAB	ChemCam Team

An index table contains one row for each product. Table 3 lists the columns of the index table.

Table 3. Index Table Contents

Column Name	Description	Source
VOLUME_ID	PDS archive volume on which a data product is stored	Specified at time of index table creation
PATH_NAME	Directory path to data product, relative to volume root	Determined by location of PDS label file
FILE_NAME	Name of *label* file for data product	PDS label file name
PRODUCT_ID	Unique identifier for data product	PDS label
PRODUCT_VERSION_ID	Version number of data product	PDS label
PRODUCT_TYPE	Type of RDR product, from Table 1	PDS label
PRODUCT_CREATION_TIME	Date and time of data product creation	PDS label
START_TIME	UTC date and time of start of observation	PDS label
STOP_TIME	UTC date and time of end of observation	PDS label
SPACECRAFT_CLOCK_START_COUNT	Spacecraft clock count at start of observation	PDS label
SPACECRAFT_CLOCK_STOP_COUNT	Spacecraft clock count at end of	PDS label

	observation	
RELEASE_ID	ID of scheduled data release to PDS in which this product was included	PDS label
PLANET_DAY_NUMBER	Number of days since landing (i.e., sol number)	PDS label

2.4 Document Directory Contents

The Document Directory contains documentation to help the user understand and use the archive data. The following files are contained in the Document Directory.

File Name	File Contents	File Provided By
DOCINFO.TXT	A description of the contents of this directory	Geosciences Node
GEOMETRIC_CM.TXT	A description of the MSL geometric camera model	OPGS
MSL_CAMERA_SIS.PDF	The Data Product SIS as a PDF file	ChemCam Team
MSL_CAMERA_SIS.LBL	PDS label that describes the CAMERA_SIS.PDF	Geosciences Node
CHEMCAM_ARCHSIS.PDF	The Archive Volume SIS (this document) as a PDF file	ChemCam Team
CHEMCAM_ARCHSIS.LBL	PDS label that describes the CHEMCAM_ARCHSIS.PDF.	Geosciences Node
MSL_LDD.FUL, MSL_LDD.LBL	MSL Local Data Dictionary and its label file	Geosciences Node
PDSDD.FUL, PDSDD.LBL	Planetary Data System Data Dictionary and its label file	Geosciences Node
MSL_CCAM_OBS,CSV MSL_CCAM_OBS.LBL	Summary information for ChemCam observations in a comma-delimited ASCII table, and its label file	ChemCam Team

2.5 Catalog Directory Contents

The files in the Catalog Directory provide a top-level understanding of the mission, spacecraft, instruments, and data sets. The files in this directory are coordinated with the PDS data engineer, who is responsible for loading them into the PDS catalog. The following files are found in the Catalog Directory.

File Name	File Contents	File Provided By
CATINFO.TXT	A description of the contents of this directory	Geosciences Node
CCAM_RMI_RDR_DS.CAT, CCAM_LIBS_RDR_DS.CAT	Data set information for the PDS catalog	ChemCam Team
CCAM_LIBS_INST.CAT	LIBS instrument information for the PDS catalog	ChemCam Team
CCAM_RMI_INST.CAT	RMI instrument information for the PDS catalog	ChemCam Team
CCAM_PERSON.CAT	Personnel information for the PDS catalog (Team and PDS personnel responsible for generating the archive)	ChemCam Team
CCAM_REF.CAT	References mentioned in other CC*.CAT files	ChemCam Team
MSL_INSTHOST.CAT	Instrument host (i.e., spacecraft) information for the PDS catalog	MSL Project
MSL_MISSION.CAT	Mission information for the PDS catalog	MSL Project

MSL_REF.CAT	References mentioned in other MSL*.CAT files	MSL Project

2.6 Calib Directory Contents

The Calib Directory contains calibration files used to process the data products, or calibration data needed to use the data products. The following files are contained in the Calib Directory.

File Name	File Contents	File Provided By
CALINFO.TXT	A description of the contents of this directory	Geosciences Node
MSL_CCAM_LIBS_CALIB.CSV	ChemCam Pre-Flight Calibration Data.	ChemCam Team
MSL_CCAM_LIBS_CALIB.LBL	PDS label describing MSL_CCAM_LIBS_CALIB.CSV	ChemCam Team
CLEANROOM	Subdirectory of calibration data from the cleanroom calibration environment. Files in this directory have names in the format YYYY_MM_DD_hhmmss_CCS.csv [.lbl], indicating the date the data were collected.	ChemCam Team

2.7 Browse Directory Contents

The Browse Directory contains reduced-size, easily viewed versions of data products to be used to help identify products of interest. Names of browse products follow the scheme QCHEMCAM_nnnnn_inst_sclk.JPG, where nnnnn is the sol number, type is LIBS1 or LIBS2, and sclk is the spacecraft clock start time. Browse products for a given ChemCam product may be identified by matching the spacecraft clock times in the file names.

The following files are contained in the Browse Directory.

File Name	File Contents	File Provided By
BROWINFO.TXT	A description of the contents of this directory	Geosciences Node
SOLnnnn	Subdirectory of browse products acquired on sol nnnn	
QCHEMCAM_nnnnn_type_sclk.JPG	Browse version of data acquired at time sclk	ChemCam Team
QCHEMCAM_nnnnn_type_sclk.LBL	PDS label for QCHEMCAM_nnnnn_type_sclk.JPG	ChemCam Team

2.8 Extras Directory Contents

The Extras Directory contains documentation, utility programs, or other materials that the user may find helpful, but that are beyond the scope of the required elements of the archive. The contents of this directory are exempt from PDS requirements for labeling, etc. The Extras Directory is optional; it is intended for "value-added" material, handy to have but not crucial for understanding the data. (If it's crucial, it is in one of the standard directories.) The following files are contained in the Extras Directory.

File Name	File Contents	File Provided By
EXTRINFO.TXT	A description of the contents of this directory	Geosciences Node
MOC_####_####.CSV	Concatenated MOC files for sols in each PDS release	ChemCam Team
MOC_####_####.LBL	PDS label for each MOC_####_####.CSV	ChemCam Team

DN2ENGUNITS_LUT.PDF   Lookup tables for converting ChemCam data numbers (DN)   ChemCam Team
                       to physical values

3   Archive Volume Format

This section describes the format of ChemCam RDR Archive Volume. Data that comprise the Archive are formatted in accordance with Planetary Data System specifications [Applicable Documents 4 and 5].

3.1 File Formats

This section describes file formats for the kinds of files contained on the Archive Volume.

3.1.1 Document File Format

Document files with the .TXT suffix exist in all directories except Data. They are ASCII files which may have embedded PDS labels. Lines in a .TXT file end with a carriage return character (ASCII 13) and a line feed character (ASCII 10). This allows the files to be readable under various operating systems.

Documents in the Document directory may contain formatting and figures that cannot be rendered as ASCII text. These documents are given as PDF (Portable Document Format) files, which may be viewed with free PDF reader software. To comply with the PDS requirement that a text copy of the documentation be archived, an HTML version of each PDF file is also provided.

3.1.2 Tabular File Format

Tabular files (.TAB suffix) exist in the Index directory. Tabular files are ASCII files formatted for direct reading into many database management systems on various computers. All fields are separated by spaces and/or commas, and character fields are enclosed in double quotation marks ("). (Character fields are padded with spaces to keep quotation marks in the same columns of successive records.) Character fields are left justified, and numeric fields are right justified. The "start byte" and "bytes" values listed in the labels do not include the commas between fields or the quotation marks surrounding character fields. The records are of fixed length, and the last two bytes of each record contain the ASCII carriage return (ASCII 13) and line feed (ASCII 10) characters. This allows a table to be treated as a fixed length record file on computers that support this file type and as a text file with embedded line delimiters on those that don't.

3.1.3 CSV File Format

ChemCam LIBS RDR data products are stored as ASCII text comma-separated-value (CSV) files. Fields are delimited by commas and may vary in length from one record to the next. Two consecutive commas indicate there is no data value for a field. Character fields are enclosed in double quotation marks. As in table files, each record is terminated with ASCII carriage return and line feed characters.

3.1.4 Image File Format

ChemCam RMI RDR data products are stored as binary image files. The accompanying PDS label gives information about the image size, pixel size, data type, etc., necessary for reading and displaying the image. See Applicable Document 3 for details.

ChemCam RDR image file names may have the extension ".IMG" or ".TIF". In either case the images are described by detached PDS labels. "IMG" is the standard file name extension for PDS images. "TIF" indicates a TIFF (Tagged Image File Format) file, a popular image file format that can be read by many image processing programs.

3.1.5 PDS Label Format

All image, tabular, and CSV files in the archive are described by detached PDS labels. A detached label has the same name as the data file it describes, but with the extension .LBL. For examples of PDS labels for each type of ChemCam data product, see the Camera EDR/RDR Data Product SIS [Applicable Document 3].

A PDS label consists of lines of ASCII text in the form of keyword = value statements that provide descriptive information about the data file. The label is intended to be readable both by humans and by software. Details of the syntax and semantics of PDS labels can be found in the PDS Standards Reference (Applicable Document 5), and definitions of the keywords used in the label can be found by using the PDS Data Dictionary Lookup web service at http://pds.nasa.gov/tools/ddlookup/data_dictionary_lookup.cfm.

Lines of text in labels end with a carriage return character and a line feed character. This allows the files to be read under various operating systems.

Image files may also contain other labels embedded at the beginning of the file. These labels are described in the Camera EDR/RDR Data Product SIS [Applicable Document 3]. In particular, an image file may contain an embedded label that looks very similar to the detached PDS label, except that it begins with the keyword ODL_VERSION_ID instead of PDS_VERSION_ID. This is the operations version of the image label, generated at the time of downlink for internal mission use. It may contain a few keywords that are different from the detached PDS label. The operations label may not comply with PDS standards.

3.1.6 Catalog File Format

Catalog files (extension .CAT) exist in the Root and Catalog directories. Like PDS labels, they are made up of keyword = value statements that are software-readable, but most of their content is descriptive text intended to be read by the user. They contain descriptions of the data set, instrument, spacecraft, and mission, as well as personnel contact information and references to published literature. They are called catalog files because they are loaded into the PDS online catalog to make the information available to users searching for data.

3.1.7 Science Data File Formats

ChemCam RMI RDR data products are images with detached PDS labels. ChemCam LIBS RDR data products are CSV files with detached PDS labels.

For more information about the format and content of the data products, see the Camera EDR/RDR Data Product SIS [Applicable Document 3].

4   Archive Volume Generation

4.1   Data Transfer and Validation Methods

ChemCam LIBS RDR data products are generated by the ChemCam Science Team, either at JPL or at investigator institutions. ChemCam RMI RDR data products (images) are generated by the MSL Operations Product Generation Subsystem (OPGS). The RDR data products are prepared from calibrated and transformed ChemCam EDR (raw) data products according to the Camera EDR/RDR Data Product SIS [3]. The ChemCam Team delivers data products to the Geosciences Node of the PDS, along with the other components of the archive for which the team is responsible, according to the archive delivery schedule in the MSL Archive Plan [2] and agreement between the ChemCam Team and the Geosciences Node [6]. Deliveries occur every 90 sols.

4.2   Interface Media Characteristics

All volumes in the ChemCam RDR Archive conform to ISO 9660 standards as specified in the PDS Standards Reference [5].

4.3   Backup

The PDS Geosciences Node maintains backup copies of MSL ChemCam RDR archives according to its general policy for backups, including off-site storage belonging to the Node and long-term storage at the National Space Science Data Center.

4.4   Revisions

It is expected that ChemCam RDR data products may be revised during the course of the mission due to improvements in the understanding of the instrument, better calibration information, etc. When the ChemCam Team delivers revised versions of ChemCam RDR data products to the PDS, the team must identify the revisions at or before the time of delivery. Labels for revised products must have a later value for the PRODUCT_CREATION_TIME keyword and an updated value for the PRODUCT_VERSION_ID keyword. They may also have a different value in the VERSION field of the file name as described in section 2.2, Data Directory Contents and Naming. The reasons for the revisions must be documented in ERRATA.TXT.

Previous versions of revised products are handled according to PDS policy for "draft" data sets, as the ChemCam RDR archive is considered a work in progress until the final delivery. If revisions consist of minor changes to labels or corrections of typographical errors, the previous versions are not retained. If the revisions are done to correct serious errors in previously released products such that the older products are unfit for use, then the older products are not retained. In the typical case of revisions done as part of routine updates in calibration, etc., the Geosciences Node retains previous versions, making them accessible to users, for six months.

4.5   Labeling and Identification

The ChemCam RDR Archive consists of three data sets, all stored on the same archive volume.

PDS Data Set ID	PDS Volume ID
MSL-M-CHEMCAM-RMI-5-RDR-V1.0	MSLCCM_1XXX

MSL-M-CHEMCAM-LIBS-4/5-RDR-V1.0	MSLCCM_1XXX

5 Support Staff and Cognizant Persons

The ChemCam Principal Investigator is Roger Wiens, Los Alamos National Laboratory. The ChemCam Archive Representatives are ChemCam Science Team members Bruce Barraclough and Dot Delapp, Los Alamos National Laboratory. The MSL Deputy Project Scientist, who is responsible for coordinating MSL archiving, is Joy Crisp, Jet Propulsion Laboratory.

The PDS Geosciences Node Archivist for the ChemCam RDR Archive is Susan Slavney, Washington University, St. Louis; she is also the MSL Mission Lead for PDS. The PDS Geosciences Node Manager is Raymond Arvidson, Washington University. The PDS Data Engineer for MSL is Betty Sword, PDS Engineering Node, Jet Propulsion Laboratory.