1.0 STANDARD OPERATING PROCEDURE (SOP): IMAGING DATA TRANSFER TO JPL

Document Version 7.4 August 19, 2024

Overview

This document describes the standard operating procedures for transfer of in vivo, human imaging data to the Jet Propulsion Laboratory (JPL) for the EDRN "Validation of Molecular Biomarkers for the Early Detection of Lung Cancer in the setting of Indeterminate Pulmonary Nodules" (Lung Team Project #2 – LTP2) project.

These data include computed tomography (CT) and positron emission spectroscopy (PET). Other modalities may be considered through extension of this SOP. The scope of this SOP only pertains to the transfer of images and security of data defined herein. Changes to this document must be confirmed and approved by Imaging Core Team (see below).

Imaging Core Team

The imaging core team consists of Drs. Deppen and Schabath who are part of the LTP2 project, along with Dan Crichton and Heather Kincaid at JPL.

Image Transfer and Repository

Study sites will transfer **de-identified** data to NASA Jet Propulsion Laboratory (JPL) for archiving in DICOM format. Prior to transferring data to JPL an anonymized **Image Event ID** must be generated in VSIMS by entering the Follow-up Image Identifier form in VSIMS Submit Data under Follow-up. The Image Event ID must be used in the LTP2 Image file names to label the CT/PET data prior to upload. The Image Event ID needs to be shared with whoever, at each site, is de-identifying and naming the imaging files for use in this study.

De-identification process:

Below is a list of Medidata to strip and metadata to retain in the DICOM header.

DICOM Tag	Attribute Name	Notes	
(0008,0005)	Specific Character Set	Example: ISO_IR 100	Retain
	Patient Name	If included	**Replace with VSIMS Image Event ID#. DO NOT include the Participant ID#**
(0010,0030)	Patient's Birth Date	Remove or use a reference date	Remove/Alter
(0010,1000)	Other Patient IDs		Remove
(0010,1001)	Other Patient Names		Remove
(0010,0020)	Patient ID		**Replace with VSIMS Image Event ID#. DO NOT include the Participant ID#**
	Medical Record #	If included	**Replace with VSIMS Image Event ID#. DO NOT include the Participant ID#**
(0008,0008)	Image Type	Example: ORIGINAL,PRIMARY,AXIAL	Retain
(0008,0016)	SOP Class UID	Example: 1.2.840.10008.5.1.4.1.1.2	Retain

DICOM Tag	Attribute Name	Notes	
(0008,0018)	SOP Instance UID	Example: 1.2.392.200036.9116.2.6.1.3268.2051314068.1558478927.941955	Retain
(0008,0020)	Study Date	YYYYMMDD	Retain
(0008,0021)	Series Date	YYYYMMDD	Retain
(0008,0022)	Acquisition Date	YYYYMMDD	Retain
(0008,0023)	Content Date	YYYYMMDD	Retain
(0008,0030)	Study Time	Example: 074654.000	Retain
(0008,0031)	Series Time	Example: 074757.337	Retain
(0008,0032)	Acquisition Time	Example: 074837.600	Retain
(0008,0033)	Content Time	Example: 074837.841	Retain
(0008,0050)	**Accession Number**	Example: **3268852**	**Replace with VSIMS Image Event ID#. DO NOT include the Participant ID#**
(0008,0060)	Modality	Example: CT	Retain
(0008,0070)	Manufacturer	Example: TOSHIBA	Retain
(0008,0090)	Referring Physician's Name	Example: Unspecified	Replace name with Unspecified
(0008,1010)	Station Name	Example: AQOne	Retain
(0008,1030)	Study Description	Example: CT CHEST LOW DOSE	Retain
(0008,1032)	Procedure Code Sequence	Example: [(0008, 0100) Code Value SH: 'CTCHLODOSE' (0008, 0102) Coding Scheme Designator SH: 'LOCAL' (0008, 0104) Code Meaning LO: 'CT CHEST LO DOSE']	Retain
(0008,103E)	Series Description	Example: Axial Body Std. Axial Non Contrast	Retain
(0008,1040)	Institutional Department Name	Example: CT	Retain
(0008,1070)	Operators Name	Example: LAS	Retain
(0008,1090)	Manufacturer Model Name	Example: Aquilion ONE	Retain
(0008,1110)	Referenced Study Sequence	Example: [(0008, 1150) Referenced SOP Class UID UI: Detached Study Management SOP Class (0008, 1155) Referenced SOP Instance UID UI: 1.2.840.113745.101000.1231000.43590.4783.4390843]	Retain
(0010,0040)	Patient Sex	N/A	Retain
(0010,1010)	Patient Age	Example: 056Y	Retain
(0010,1005)	Patient Birth Name		Remove
(0010,2160)	Ethnic Group		Remove
(0010,2180)	Occupation		Remove

DICOM Tag	Attribute Name	Notes	
(0010,1060)	Patient Mother's Birth Name		Remove
(0010,2154)	Patient's Telephone Numbers		Remove
(0010,1040)	Patient's Address		Remove
(0018,0015)	Body Part Examined	Example: CHEST	Retain
(0018,0022)	Scan Options	Example: HELICAL_CT	Retain
(0018,0050)	Slice Thickness	Example: 5.0	Retain
(0018,0060)	KVP	Example: 120	Retain
(0018,0090)	Data Collection Diameter	Example: 400.00	Retain
(0018,1000)	Device Serial Number	Example: 8DA1582002	Retain
(0018,1020)	Software Versions	Example: V7.03ER919	Retain
(0018,1030)	Protocol Name	Example: Low Dose Lung	Retain
(0018,1100)	Reconstruction Diameter	Example: 379.687	Retain
(0018,1120)	Gantry Detector Tilt	Example: +0.0	Retain
(0018,1130)	Table Height	Example: +115.00	Retain
(0018,1140)	Rotation Direction	Example: CW	Retain
(0018,1150)	Exposure Time	Example: 275	Retain
(0018,1151)	X Ray Tube Current	Example: 27	Retain
(0018,1152)	Exposure	Example: 7	Retain
(0018,1160)	Filter Type	Example: LARGE	Retain
(0018,1170)	Generator Power	Example: 3	Retain
(0018,1190)	Focal Spots	Example: 0.9,0.8	Retain
(0018,1210)	Convolution Kernel	Example: FC18-H	Retain
(0018,5100)	Patient Position	Example: FFS	Retain
(0018,9302)	Acquisition Type	Example: SPIRAL	Retain
(0018,9305)	Revolution Time	Example: 0.275	Retain
(0018,9306)	Single Collimation Width	Example: 0.5	Retain
(0018,9307)	Total Collimation Width	Example: 40.0	Retain
(0018,9310)	Table Feed Per Rotation	Example: 32.5	Retain
(0018,9311)	Spiral Pitch Factor	Example: 0.813	Retain
(0018,9313)	Reconstruction Target Center Patient	Example: 0.0,0.0,1938.5	Retain
(0018,9323)	Exposure Modulation Type	Example: 3D	Retain
(0018,9324)	Estimated Dose Saving	Example: 39.4	Retain
(0018,9327)	Table Position	Example: 0.0	Retain

DICOM Tag	Attribute Name	Notes	
(0018,9331)	Fluoroscopy Flag	Example: NO	Retain
(0018,9345)	C T D Ivol	Example: 0.9	Retain
(0020,000D)	Study Instance U ID	Example: 1.2.392.200036.9116.2.6.1.3268.2051314068.1558478805.330002	Retain
(0020,000E)	Series Instance U ID	Example: 1.2.392.200036.9116.2.6.1.3268.2051314068.1558478927.941051	Retain
(0020,0010)	**Study ID**	**Example: 3268852**	**Replace with VSIMS Image Event ID#. Do NOT include the Participant ID#**
(0020,0011)	Series Number	3	Retain
(0020,0012)	Acquisition Number	Example: 3	Retain
(0020,0013)	Instance Number	Example: 1	Retain
(0020,0020)	Patient Orientation	Example: L,P	Retain
(0020,0032)	Image Position Patient	Example: -189.4729,-189.4729,1938.50	Retain
(0020,0037)	Image Orientation Patient	Example: 1.00000,0.00000,0.00000,0.00000,1.00000,0.00000	Retain
(0020,0052)	Frame Of Reference U ID	Example: 1.2.392.200036.9116.2.6.1.3268.2051314068.1558478816.341703	Retain
(0020,1040)	Position Reference Indicator	Example: N/A	Retain
(0020,1041)	Slice Location	Example: +0.00	Retain
(0020,4000)	Image Comments	Example: Non \Contrast	Retain
(0028,0002)	Samples Per Pixel	Example: 1	Retain
(0028,0004)	Photometric Interpretation	Example: MONOCHROME2	Retain
(0028,0010)	Rows	Example: 512	Retain
(0028,0011)	Columns	Example: 512	Retain
(0028,0030)	Pixel Spacing	Example: 0.741,0.741	Retain
(0028,0100)	Bits Allocated	Example: 16	Retain
(0028,0101)	Bits Stored	Example: 16	Retain
(0028,0102)	High Bit	Example: 15	Retain
(0028,0103)	Pixel Representation	Example: 1	Retain
(0028,1050)	Window Center	Example: 40	Retain
(0028,1051)	Window Width	Example: 400	Retain
(0028,1052)	Rescale Intercept	Example: 0	Retain
(0028,1053)	Rescale Slope	Example: 1	Retain

DICOM Tag	Attribute Name	Notes	
(0032,1060)	Requested Procedure Description	Example: CT CHEST LOW DOSE	Retain
(0032,1064)	Requested Procedure Code Sequence	Example: [(0008, 0100) Code Value SH: 'CTCHLODOSE' (0008, 0102) Coding Scheme Designator SH: 'LOCAL' (0008, 0104) Code Meaning LO: 'CT CHEST LOW DOSE']	Retain
(0040,0002)	Scheduled Procedure Step Start Date	YYYYMMDD	Retain
(0040,0003)	Scheduled Procedure Step Start Time	Example: 080000	Retain
(0040,0004)	Scheduled Procedure Step End Date	YYYYMMDD	Retain
(0040,0005)	Scheduled Procedure Step End Time	Example: 083000.000	Retain
(0040,0007)	Scheduled Procedure Step Description	Example: CT CHEST LOW DOSE	Retain
(0040,0008)	Scheduled Protocol Code Sequence	Example: [(0008, 0100) Code Value SH: 'RISIC' (0008, 0102) Coding Scheme Designator SH: 'RISIC' (0008, 0104) Code Meaning LO: 'RISIC']	Retain
(0040,0009)	Scheduled Procedure Step ID	Example: CTCHLODOSE	Retain
(0040,0244)	Performed Procedure Step Start Date	YYYYMMDD	Retain
(0040,0245)	Performed Procedure Step Start Time	Example: 074654.000	Retain
(0040,0253)	Performed Procedure Step ID	Example: 41877	Retain
(0040,0260)	Performed Protocol Code Sequence	Example: [(0008, 0100) Code Value SH: 'RISIC' (0008, 0102) Coding Scheme Designator SH: 'RISIC' (0008, 0104) Code Meaning LO: 'RISIC']	Retain
(0040,0275)	Request Attributes Sequence	Example: [(0040, 1001) Requested Procedure ID SH: '3268852']	Retain
(0040,2017)	Filler Order Number Imaging Service Request	Example: 3268852	**Replace with VSIMS Image Event ID#. Do NOT include the Participant ID#**
(0088,0200)	Icon Image Sequence	Example: [(0028, 0002) Samples per Pixel US: 1 (0028, 0004) Photometric Interpretation CS: 'MONOCHROME2' (0028, 0010) Rows US: 128 (0028, 0011) Columns US: 128 (0028, 0100) Bits Allocated US: 8 (0028, 0101) Bits Stored US: 8 (0028, 0102) High Bit US: 7 (0028, 0103) Pixel Representation US: 0 (7fe0, 0010) Pixel Data OW: Array of 16384 elements]	Retain
(0018,0015)	File Type	Example:dicom	Retain
(0018,0022)	File Size	Example:549.8 kB	Retain

DICOM Tag	Attribute Name	Notes	
(0018,0050)	File Download Id	Example: N/A	**Replace with VSIMS Image Event ID#. Do NOT include the Participant ID#**

Example File Naming Convention for a given CT or PET:

IMG#1_EventIdentifier.dcm

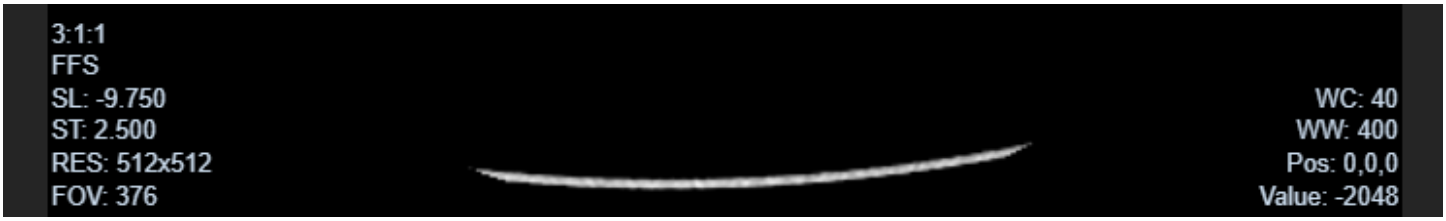
IMG#2_EventIdentifier.dcm

Etc. All Images from an Imaging Test done on a given date.

EXAMPLE: Header:

A. Replace with VSIMS Image Event Identifier. Do not include the Participant ID
B. Replace with VSIMS Image Event Identifier. Do not include the Participant ID
C. Replace with random date
D. Ensure it says Anaonymous
E. Retain Date and Time Stamp
F. Retain Image Type

EXAMPLE: Footer: Retain All



Data will be deposited into the EDRN Cancer Biomarker Data Commons (LabCAS), a web-enabled environment that allows users to publish, share, search and download a wide variety of biomedical datasets. In LabCAS, data is organized according to the following logical hierarchy: Collections: broad sets of related data from the same study, the same analysis, or the same project; Datasets: different sets of related files within the same collection; and Files: all the files in a given dataset. Metadata will be associated with the data that will include the link to the de-identified clinical data stored in VSIMS. Sites will use IBM Aspera to transfer files to JPL.

Step 1. Documentation of authorization for data transfer.

Each recruiting site will send proof of IRB approval to the DMCC allowing images to be transferred to JPL. The DMCC will ensure that JPL has IRB approval for warehousing the images. If the recruiting site requires a Data Use Agreement between their institution and JPL, then the recruiting institute is responsible for initiating the DUA. Each recruiting site will determine the following (1) the number of subjects and files to be transferred, (2) the nature of the transfer (prospective/ongoing collection), (3) the process by which data will be de-identified, (4) a technical contact who will facilitate data transfer and (5) a data access policy for de-identified imaging data.

Step 2. Definition of transfer protocol and de-identification method

De-identified images will be sent to JPL via IBM Aspera Connect, a secure ,fast and reliable data transfer transfer tool.

Data transfer: For prospective studies, each situation requires careful consideration to minimize disruption to the existing workflow. The Imaging Core will discuss with technical staff to establish a standardized method that is both flexible and seamless with existing workflow and data transfer procedures.

Data de-identification: JPL does not store identified imaging data on its imaging servers. The DMCC maintains a link that links imaging identifiers, the Imaging Event ID, with patient identifiers and other common data elements (CDEs) defined by the study all in a de-identified manner. JPL requires that each recruiting site de-identifies all images prior to data transfer to meet their IRB requirements.

Step 3. Transfer, review, and de-identification of data

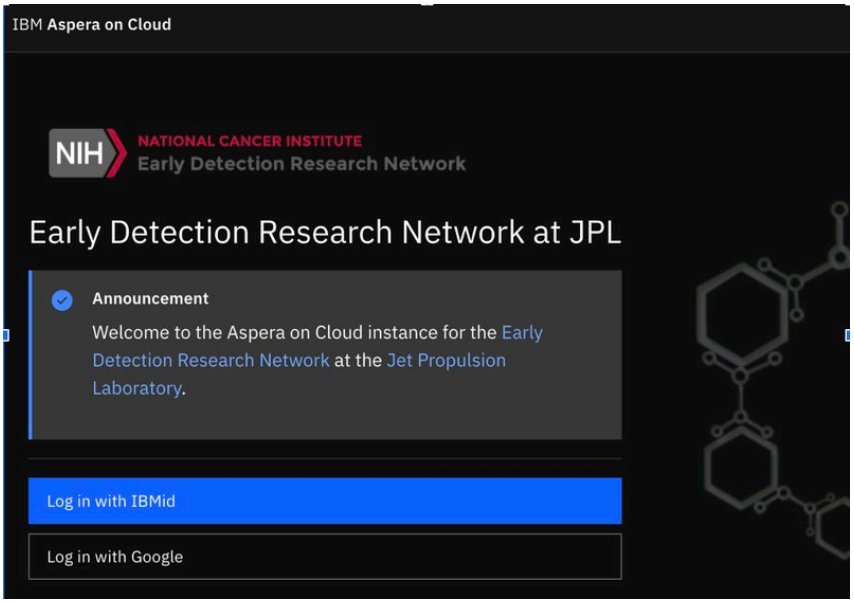
Setting up Data Transfer (Getting an Aspera Account)

To upload data into LabCAS, you'll use IBM Aspera Connect.

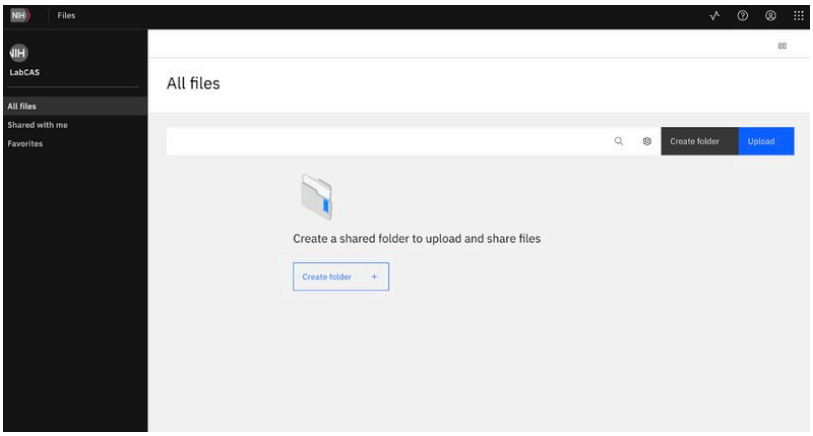
1. Contact the JPL Informatics Center by email at ic-data@jpl.nasa.gov: Include the following:
    a. Your name and Institution
    b. The project name for which you will be uploading data, such as the "EDRN Lung Team Project 2"
2. Upon receipt of your email, the JPL team will send you an invitation to join the Aspera on Cloud instance designated for EDRN.
3. Accept the Invitation: Look out for the email invitation in your inbox, and click the "Accept" button within the message. This action will direct you to the EDRN Aspera login page.

Uploading your data

1. Navigate to EDRN Aspera by visiting → https://edrn-labcas.ibmaspera.com/:



2. Sign in using your Aspera credentials.



3. Navigate to the files on your computer and drag and drop

   OR

   Click the 🔍 to find the files

4. Click Upload

Note: The first time, a small helper application, "Aspera Connect", should be downloaded and installed.

5. **Notify the JPL Informatics Center** that your data transfer is complete by emailing ic-data@jpl.nasa.gov
   - Include your VSIMS username so we can provide you with permissions to view your sites images in LabCAS.

Receiving, reviewing and publishing data

6. The JPL team will review and ingest your data to the EDRN Cancer Biomarkers Data Commons (LabCAS).

They will provide access to the data submission team for review. The data that is received will be compared to the Event IDs expected. Any and all error(s) that may happen will be reviewed and documented (with special consideration that no identifying data is stored in such documents). JPL will communicate with sites if there are any data transfer errors as soon as pragmatic to ensure efficient and accurate data transfer.

JPL will provide access to the images

Review your uploaded images in LabCAS. Use your EDRN/VSIMS login to LabCAS→ https://edrn-labcas.jpl.nasa.gov/

Contact us if you have any questions at ic-data@jpl.nasa.gov

Review your images in LabCAS

7. JPL will notify you once your images are in LabCAS.
8. Login to EDRN LabCAS at (https://edrn-labcas.jpl.nasa.gov/labcas-ui/) using your EDRN credentials and review your images for completeness.

Step 4. Data Access

After quality assurance, the imaging data will be accessible to the appropriate groups within LabCAS. Access will be granted via web interfaces and controlled via passwords. Initially, only the DMCC statisticians will have

access to the images for analysis. Once the DMCC develops a process with the investigators to ensure blinding, images may be shared upon approval.

Data Security

Clinical and Specimen Data

All Clinical Data is stored in VSIMS using a de-identified patient identifier. The DMCC and all servers are physically secured behind a card-key access area. Fred Hutchinson Cancer Research Center (FHCRC) staff with authorized access to the DMCC may enter the physical area. All materials and data are stored inside locked offices and server rooms. All DMCC servers are secured in a locked area with extremely limited key distribution. All servers require username and password logins and file permissions are granted on an as needed basis. All DMCC computers have a mandatory screensaver password that is unique for each user.

VSIMS has three levels of security. The first level is a login system that requires a username and password. The second level is the assignment of protocol access to a specified user. For example, if a user is authorized to access a single specified protocol, but VSIMS is managing data for three protocols at that time, the user is only allowed to access that single specified protocol. The third level is the assigned user rights as described in detail below. These user rights are assigned by protocol.

Access to VSIMS requires a username and password distributed and maintained by the DMCC. The username and password are the keys to accessing VSIMS. Everyone who accesses VSIMS should keep this information as safe as the keys to the house or car. Keeping the username and password safe will protect the files in the user's pages of interest and will help guard against unauthorized use. If unscrupulous computer trespassers obtain the username or password, they can use the user's identity as a "home base" to break into the entire web site and database.

To obtain VSIMS access, one must complete an on-line VSIMS Access Application. The applicant must electronically indicate commitment to confidentiality and completion of human subjects training offered within the electronic application form. The DMCC Project Director (or assigned designee) must approve, via email, the applicant for access to VSIMS and assign user rights. (this process is documented in LTP2 MOP-Appendix 1).

Once the application is processed at the DMCC, the applicant is sent a link, verification code and username via e-mail by which to create a password for future login to the website. The user has three days in which to use the link and validation code to login and change their password. If a user does not use the link and validation code before they expire, they must re-apply.

The DMCC requires the password to be changed every six months. In addition, a user can change his/her password at any time. For security reasons, passwords for the VSIMS secure web site must be at least 8 characters long. If the user's session remains idle for 2 hours, they will be timed out and must log into the system again. Passwords or log-in information may not be shared. If a person attempts to log-in and his/her password has expired, the user is prompted to change his/her password at that time.

Acceptance of the Confidentiality Pledge and completion of the VSIMS Access Application is tracked by a database at the DMCC. A report can be generated at any time to show the VSIMS secure site users.

For security purposes, accounts that are not used for six months are deactivated and accounts that are not used for one year are deleted. Deactivation of an account will require the user to call the DMCC to reactivate it. If a user of a deleted account wants to regain access to the secure web site, he/she must complete a new Access Application. Once a year the DMCC will send an email to the Project Coordinator of each site with a list of people from his/her site who have access to the VSIMS secure site to confirm whether or not all those listed should still have access.

In addition to a username and password, to access VSIMS the user must be using a browser, which supports at least 128-bit strong data encryption. 128-bit strong encryption is recognized as the de facto standard for the secure exchange of information and is the highest internationally available level of encryption used for the exchange of data over the public Internet. Once a connection is made to the site, all communications between the user's browser and the web site are then encrypted. Further security is provided by the use of an authentication certificate provided by a major commercial Certificate Authority such as Thawte, and the DMCC's institutional firewall which blocks access to TCP/IP services not needed for accessing the secure site. The DMCC will continue to monitor the technology and policy changes that allow for continued privacy in data sharing to best serve the needs of EDRN. The DMCC must be notified immediately if a staff member that has VSIMS access no longer works on an assigned protocol so that their account can be disabled.

Imaging Data

De-identified imaging data will be loaded into LabCAS. Access will be controlled via a secure web login. Users may access the data via https://edrn-labcas.jpl.nasa.gov/ui/c upon approval. LabCAS provides a comprehensive security infrastructure including both authentication and authorization. Security services are managed using the Lightweight Directory Access Protocol (LDAP). LDAP manages both users and groups support both authentication of users and mapping of users to groups. Data is annotated into this scheme using a multi-level security architecture. This enables data to be mapped to users who have different roles including data producers, data stewards, and data users whose privileges will be granted and enforced by the system. The data producers can identify which data users can access the data. Data stewards can support annotation and capture of the data. This allows data to be shared with a different investigator or set of investigators. Data stewards can work with the data coordinating center to help appropriately annotate the data for these scenarios. Data is transferred to LabCAS using full 128-bit encryption.

Imaging Data Server Security

Amazon Web Services (AWS) secure cloud platform will be used for the storage of the images. AWS and JPL work on a shared responsibility where AWS manages the security of the cloud and JPL manages the security in the cloud. The practical application of this model is that AWS maintains underlying hardware, physical security, and Amazon tool interface security (https://aws.amazon.com/compliance/data-center/controls/) while JPL maintains the security of the virtual network and systems used by the project. In addition, JPL has certified that AWS' security controls meets the JPL security requirements as described in JPL SaaS Cyber Security Requirements (Rev 4.0) Each virtual system in the cloud, such as the EDRN sftp server, fall under a JPL security plan and meet all JPL security requirements such as regular patching, authentication restrictions, network restrictions, and logging. EDRN's AWS instances are including in the JPL IT Security Database (ITSDB) plan 220 which is certified yearly as part of JPL's C&A process.

Image Dataset Naming Convention

The VSIMS generated Imaging Event ID, as discussed previously in section *Image Transfer and Repository* will be used to label image datasets prior to transferring to the JPL image repository.