ACEMID Data Uploader – User Manual

Author: An Zhao

Australian Imaging Service (AIS)

The University of Sydney

Date: 2 June 2025

Overview

The **Australian Centre of Excellence in Melanoma Imaging & Diagnosis (ACEMID) Uploader** is a robust, automated system developed to streamline the transfer of ACEMID imaging data from a staging environment to a centralized XNAT server. This system is a critical component in the ACEMID data pipeline, ensuring that high-resolution imaging data is reliably and efficiently uploaded for further analysis and research.

Key features of the ACEMID Uploader include:

- **Real-time Monitoring**: Continuously observes the staging environment for new data.
- **Automated Upload Pipeline**: Executes scheduled uploads without manual intervention.
- **Intelligent Error Handling**: Detects and categorizes upload failures, enabling targeted reprocessing.
- **Comprehensive Logging**: Maintains detailed logs for auditing and troubleshooting.
- **Upload Verification**: Confirms the integrity and completeness of each data transfer.

Data Acquisition

The ACEMID data originates from the Canfield Vectra system, a sophisticated imaging platform that captures the entire exposed skin surface using an array of camera. This system is used in clinical settings to document and monitor skin conditions, particularly for melanoma diagnosis. Clinical staff utilize a GUI-based tool called VectraDBTool (see *figure 1*) to select and export patient data from the Vectra database.

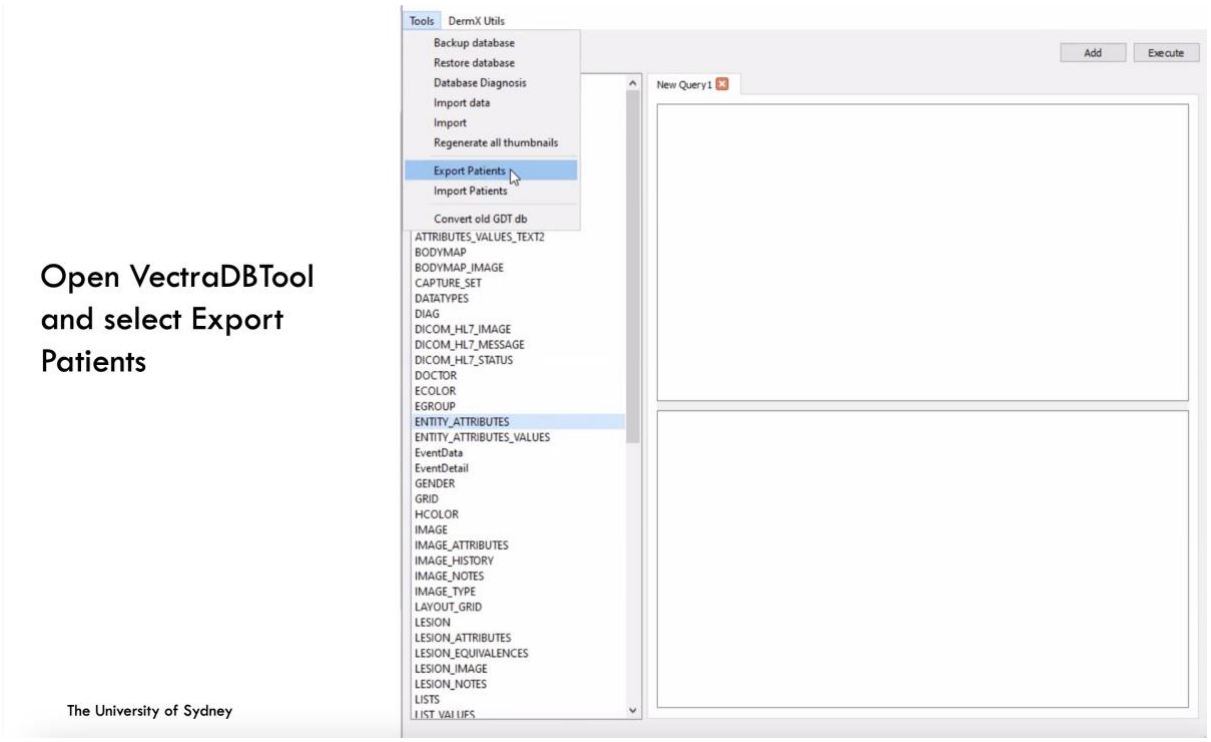

Figure 1: Export the patient ACEMID data from Canfield Vectra System

For each patient, the export process generates:

- One SQLite file containing metadata and clinical information
- One Patient folder containing high-resolution image files (see *figure 2*)

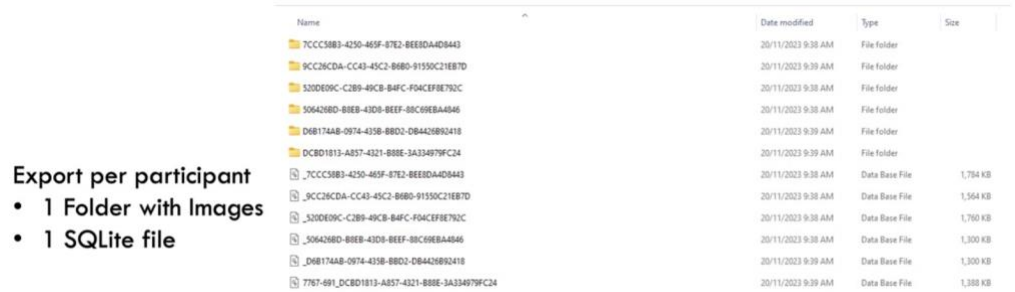

Figure 2: The Exported ACEMID data with .db SQLite file and Patient folder

Staging Server Structure

Once exported, the data is placed on a staging server with the following directory structure (see *figure 3*):

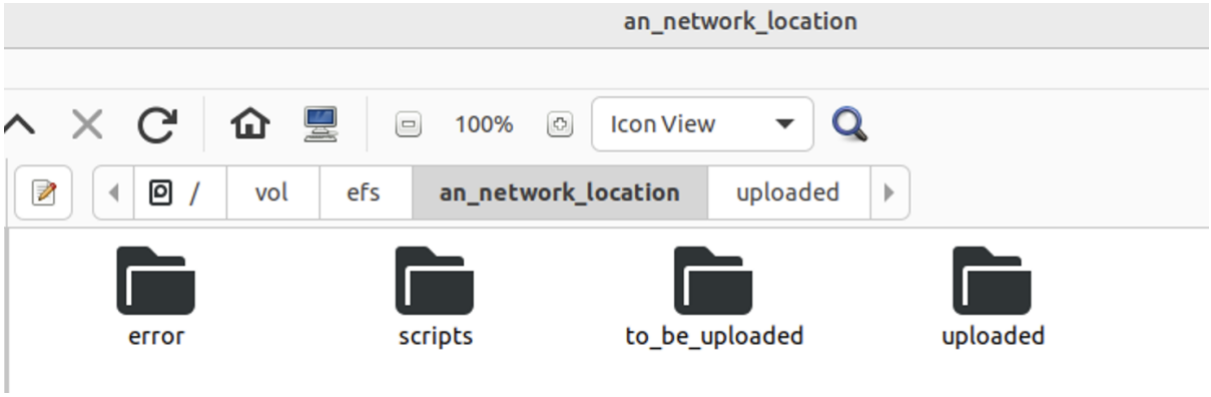

Figure 3: The folder structure of staging server for ACEMID data

- ***to_be_uploaded/***: Contains newly exported patient data awaiting upload.
- ***uploaded/:*** Stores data that has been successfully uploaded to the XNAT server.
- ***error/:*** Stores data that failed to upload due to errors.
- ***scripts/:*** The ACEMID uploader bash scripts.

Upload Process

A cron job is scheduled to execute the ACEMID uploader scripts located in the scripts/ directory. The uploader performs the following operations:

- Scans the ***to_be_uploaded/*** directory for new data.
- Initiates the upload process to the central XNAT server.
- Verifies the success of each upload.
- Moves successfully uploaded data to the ***uploaded/*** directory.
- Moves failed uploads to the ***error/*** directory.
- Another corn job to run the ACEMID uploader scripts on the ***error/*** directory.

The above process continues iteratively until the ***to_be_uploaded/*** directory is empty.

The flow chart below (see *figure 4*) outlines the steps involved in the ACEMID uploading process.

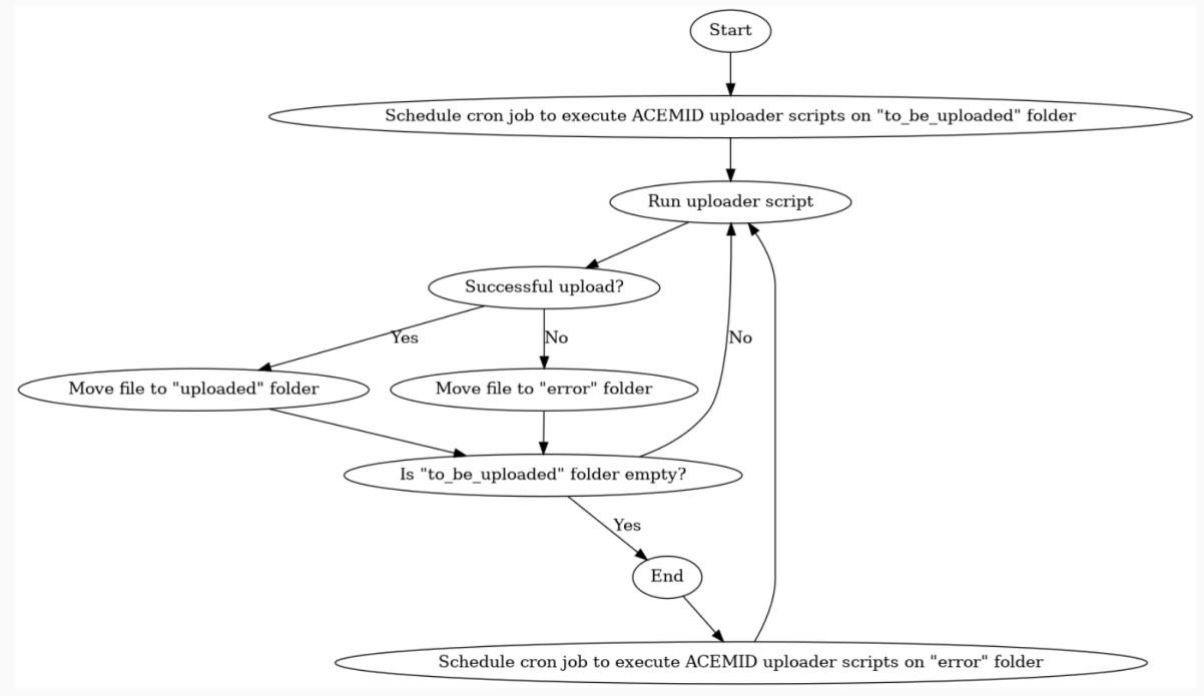

Figure 4: The flow chart of ACEMID upload process

The following pseudo (see *figure 5*) code outlines the logic of the ACEMID upload process.

```
BEGIN

// Cron Job 1: Process files in 'to_be_uploaded' folder
WHILE 'to_be_uploaded' folder is NOT empty DO
    FOR each file IN 'to_be_uploaded' folder DO
        RUN uploader_script(file)

        IF upload_successful THEN
            MOVE file TO 'uploaded' folder
        ELSE
            MOVE file TO 'error' folder
        ENDIF
    ENDFOR
ENDWHILE

// Cron Job 2: Process files in 'error' folder
WHILE 'error' folder is NOT empty DO
    FOR each file IN 'error' folder DO
        RUN uploader_script(file)

        IF upload_successful THEN
            MOVE file TO 'uploaded' folder
        ELSE
            LOG error
            RETRY upload (optional: up to N times)
            IF still unsuccessful THEN
                NOTIFY admin OR log persistent failure
            ENDIF
        ENDIF
    ENDFOR
ENDWHILE

END
```

Figure 5: The pseudo code for implementing the ACEMID upload process

Running ACEMID Uploader

Before running the uploader, ensure the following:

System Requirements

- Unix-like OS (Linux/macOS)
- Bash shell

XNAT Requirements

- Turn on the External Camera Session (xnat:xcSessionData) and External Camera Scan (xnat:xcScanData) data types (see *figure 6).*

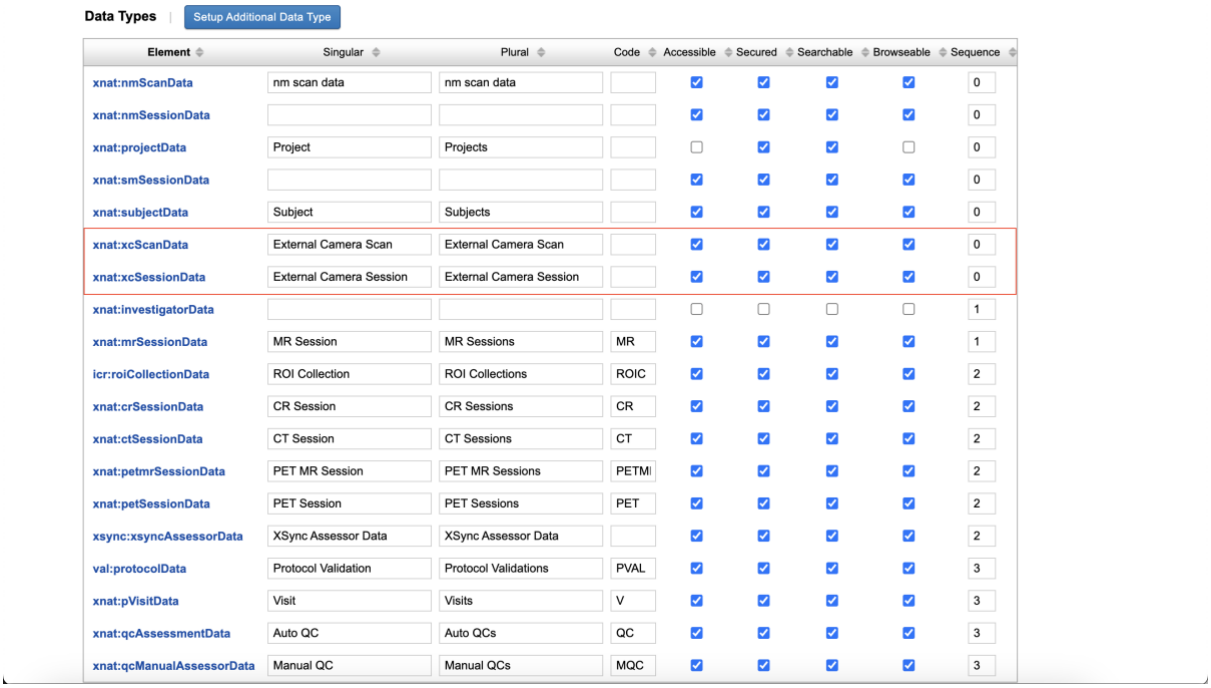

Figure 6: Turn on the xnat:xcSessionData and xnat:xcScanData type

- Suggested to Set the User Session Timeout to be 60 minutes (see *figure 7*).

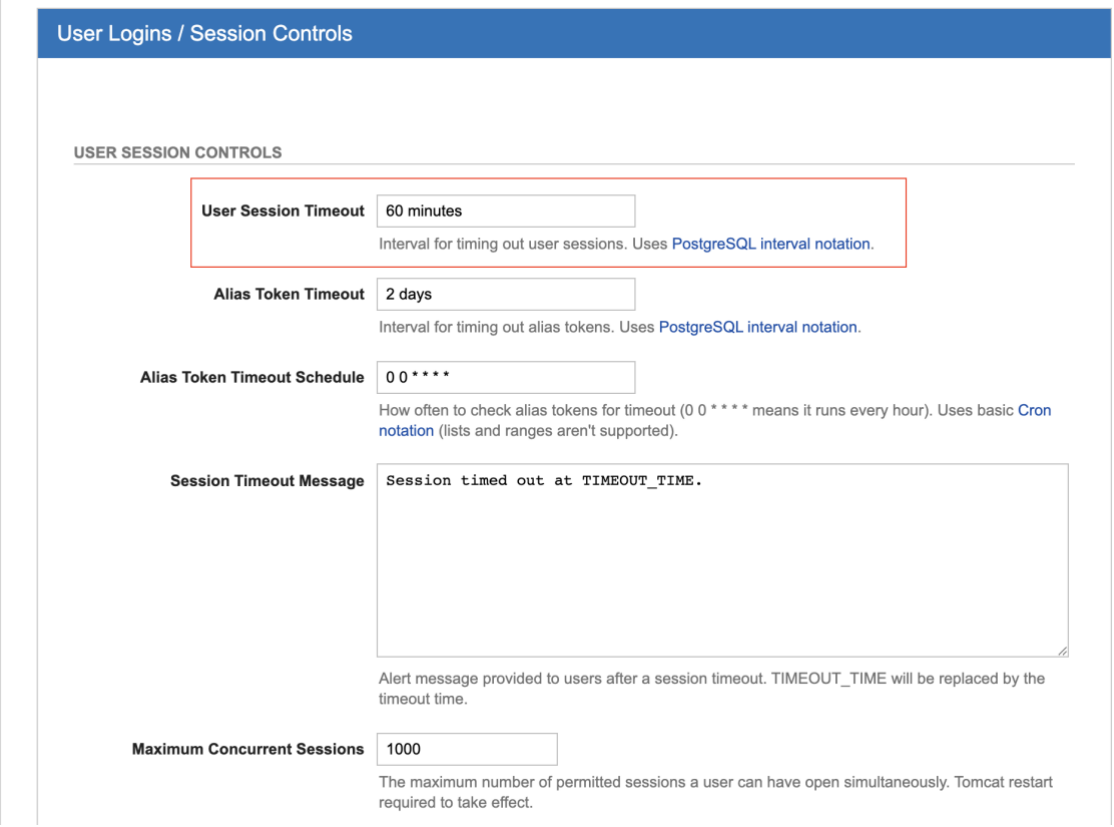

Figure 7: User Session Timeout set to 60 minutes

Install the following dependencies using your package manager. For example, on Ubuntu: run

```
sudo apt update
sudo apt install csvkit pdftotext inotify-tools
```

- `curl` (for calling XNAT Restful APIs)
- `csvkit` (for processing the csv files)
- `pdftotext` (for converting PDF files to text files)
- `enscript` (for converting text files back to PDF files)
- `inotify-tools` (for monitoring changes to files and directories in real time)

Explanations of each bash script

(1) ACEMID_uploader.sh The main ACEMID bash script to upload the cleaned Vectra exported data files to your XNAT instance using JSESSIONID.

(2) dermoscopy_data_upload.sh The Bash script to upload the dermoscopy images (mainly in jpg or png) to your XNAT instance.

(3) stage_server_monitor.sh The bash script to monitor the exported data from Vectra system to your specified network drive to detect if there is any file or folder changes in real time and then it will trigger the upload script.

(4) remove_phi_report.sh The bash script used to remove the PHI info in the pdf reports.

(5) Dockerfile The Dockerfile used to build the docker image of the above bash scripts to run on different platforms.

Setup Instructions

(1) Clone the Repository

```
git clone https://github.com/Australian-Imaging-Service/acemid-uploader.git
cd acemid-uploader
```

(2) Configure Environment Variables

Export the following variables in your shell:

```
export XNAT_URL="your_xnat_url"
export USERNAME="your_xnat_username"
export PASSWORD="your_xnat_password"
export PROJECT_ID="your_xnat_project_id"
```

You can also modify the script directly to hardcode these values if preferred.

(3) Running the script

run *./your_ACEMID_uploader_script.sh* on the terminal after you have added the execute rights to the bash script.

Conclusion

The ACEMID Uploader is a vital component of the melanoma imaging data pipeline, providing a reliable, automated mechanism for transferring clinical imaging data to a centralized repository. Its design emphasizes robustness, transparency, and fault tolerance, ensuring that valuable clinical data is preserved and made available for research and diagnostic purposes.