# SIM Data Processing

**Overview**

The system is designed to handle the upload, validation, and storage of files containing SIM records. The process involves checking whether a file has already been processed, parsing the file for validation, and storing the validated data in a database. A report is generated at the end, summarizing the processing results.

**Process Description**

1. **File Upload**
   * The system waits for the user to upload a file.
   * The file is temporarily stored in an input folder for processing.
2. **File Check for Previous Processing**
   * Once a file is uploaded, the system checks whether it has already been processed:
     + **Check in Volatile Memory (can be store in Database)**: The system checks if the file’s name or identifier exists in volatile memory or a database, indicating whether it has been processed before.
     + **If Processed**: If the file has been processed, it is deleted from the input folder to prevent reprocessing.
     + **If Not Processed**: The system proceeds to the next step for parsing and validation.
3. **File Format Structure**
   * The uploaded file is assumed to have a fixed format. The format consists of three sections:
     + **Header & Input Variables** (Rows 1–11)
     + **Records** (Data rows representing IMSIs and associated records) (small issue with the data Header sequences(IMSI is second field but in header its given as first))
   * The system will parse and validate the sections separately:
     + The **Header** provides metadata (like quantity and IMSI range) necessary for validation.
     + **Records** are the actual data entries, which are validated based on the criteria described below.
4. **Validation Process**
   * **Quantity Match**: The system verifies that the quantity defined in the **Header** matches the number of records present in the **Records** section.
     + If the count does not match, the file is marked as invalid and will not be processed further.
   * **IMSI Range Validation**: The **Header** defines the starting IMSI range, and the system will check if all IMSIs in the **Records** section fall within this valid range.
   * **Duplicate IMSI Check**: The system checks for duplicate IMSI values within the file. If a duplicate is found, the record is flagged for review.
   * **Already Existing IMSI**: The system compares each IMSI against existing IMSIs stored in the database. If the IMSI is already present in the database, the system flags the record.
5. **Processing Outcome**

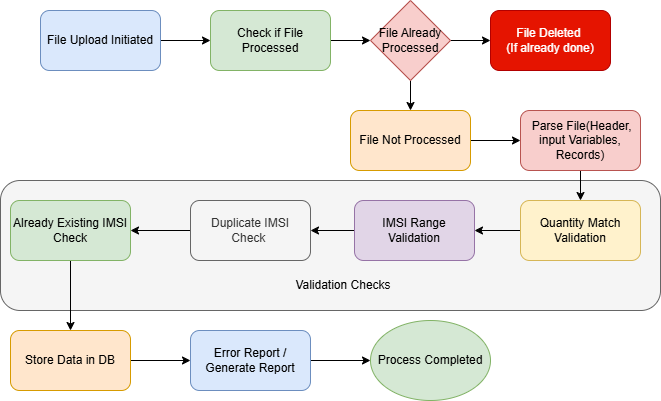
* **Validation Passed**: Once all validations (quantity, IMSI range, duplicates, and existing IMSIs) pass, the records are stored in the database.
* **Validation Failed**: If any validation fails, the file is not processed, and an error report is generated(currently validation is put only for quantity check else we are just putting the counter and giving in report summery).

1. **Report Generation**

* After processing the file, a report is generated summarizing the results of the validation:
  + Successful processing.
  + Invalid records or issues found during validation.
  + Total number of processed records.
* The report is made available for review.

**Process Flow Diagram**

Below is a high-level flow diagram representing the process:



**Explanation of the Flow Diagram:**

1. **File Upload Initiated**: The process starts when a file is uploaded.
2. **Check if File Processed**: The system checks whether the file has been processed already by querying the database or volatile memory.
3. **File Already Processed**: If the file has been processed, it is deleted from the input folder.
4. **File Not Processed**: If the file has not been processed, it proceeds to parse the content.
5. **Parse File**: The file is parsed into three sections:

* Header
* Input Variables
* Records

1. **Validation Checks**: Several validation checks are performed on the data:
   * **Quantity Match**: Ensures the number of records matches the quantity specified in the header.
   * **IMSI Range**: Verifies that the IMSI values are within the allowed range.
   * **Duplicate IMSI**: Ensures that no IMSI appears more than once in the file.
   * **Already Existing IMSI**: Checks if the IMSI already exists in the database.
2. **Validation Passed/ Error Report**: If all validations pass, the data is stored in the database.
3. **Generate Report**: Once the file is processed, a final report summarizing the processing results is generated.