# Architecture notes:

## Application objectives:

Generate GCFR script (or other tools)

Test the ETL script

## Components diagram:

Schema File Ingestion (visual interface

Version control

Persistence layer

Script Generation

Creation of script files

Testing on TeraData

User stories:

* As a database administrator I want to visualization the input schema definition

## Points to discuss :

Constant TeraData values (example: Party table names) please provide a list of those tables as used in the tabs

What is the output of the tool?

* Gcfr script to register tables (insital load)
* Incremental registration

Integration with IBM dataStage (Informatica experience):

Access to Teradata for testing (GCFR commands?)

## Tools:

Interface (streamlit)

Cursor (<https://www.cursor.com/>)

Persistence: (dataframe pickle) parquet postgres

Mermaid : design

# Design

<https://medium.com/top-python-libraries/8-powerful-python-libraries-to-create-block-diagrams-and-flowcharts-25ff04cbae00>

Block diagram and flow chart

**Code Your Diagrams: Automate Architecture with Python's Diagrams Library**

<https://dev.to/epam_india_python/code-your-diagrams-automate-architecture-with-pythons-diagrams-library-4o5o>

# Using streamlit

<https://arturlunardi.medium.com/explore-make-predictions-and-evaluate-your-ml-models-with-streamlit-and-pipelines-b6c3efeb92ff>

conda create –name strmlt

<https://medium.com/@richardhightower/streamlit-part-2-layouts-components-and-graphs-8232ef352ed1>

conda create -n strmlt python=3.12

# Building an ETL App with Streamlit

<https://medium.com/@furkan-aslan/building-an-etl-app-with-streamlit-f1334f5a7b4a>

C:\Workspaces\streamlit-tutorial\streamlit\_ETL\_tutorial.py

|  |  |
| --- | --- |
| Done | pip install streamlit |
|  | import streamlit as st import pandas as pd |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source System ID** | **Source System Name** | **Source System Alias** | **Schema** | **Loading Type** | **Source Type** |
| 1 | Core Banking | CB | CB | Daily | Oracle |
| 2 | Card Management | CMS | CMS | Daily | SQLSERVER |

[EXEC GCFR\_Register\_System('1','/','CB','Core Banking');](https://medium.com/@furkan-aslan?source=post_page---byline--f1334f5a7b4a---------------------------------------)

# Automating DBT Script Generation for Scalable Data Migration

<https://medium.com/@shakuntala.majumder/automating-dbt-script-generation-for-scalable-data-migration-10eac7cab55e>

# Workspace

C:\Workspaces

Currently testing in the folder :

C:\Workspaces\streamlit-tutorial

**To test the script creation**

C:\Workspaces\streamlit-tutorial\reading\_excel\_file.py

**The streamlit application**

C:\Workspaces\streamlit-tutorial\streamlit\_ETL\_tutorial.py