

# Project Plan Presentation

## Data Consistency and Reconciliation Tool

### The Capstone Experience

#### Team Ally

Venkata Chinmayee Mannava

Linh Nguyen

Abishek Pemmada

Tinku Sharma

Julia Sznitka

Jordan Tansingco

Department of Computer Science and Engineering

Michigan State University



*From Students...*  
*...to Professionals*

Fall 2025

# Project Sponsor Overview

- Financial company that was originally founded by GM for auto financing services.
- One of the leading financial service companies in the United States, offering services that range from insurance to loans.
- Company has a "Do It Right" mission to the people first.



# Project Functional Specifications

- Countless hours spent on manual data management, analysis, and reconciliation of data
- Business users wish to have a faster, efficient way
- Web-based tool streamlining this process:
  - Connecting to numerous data sources while also allowing users to log in and upload their data sources
  - Run consistency checks utilizing ML tools
  - Generate dashboards on a user-friendly interface highlighting the summary of the reports and discrepancies
  - Provides the user with the access and the option to reconcile any of the issues manually
- Ensure consistency and accuracy across sources, improving the confidence in reporting and decision-making



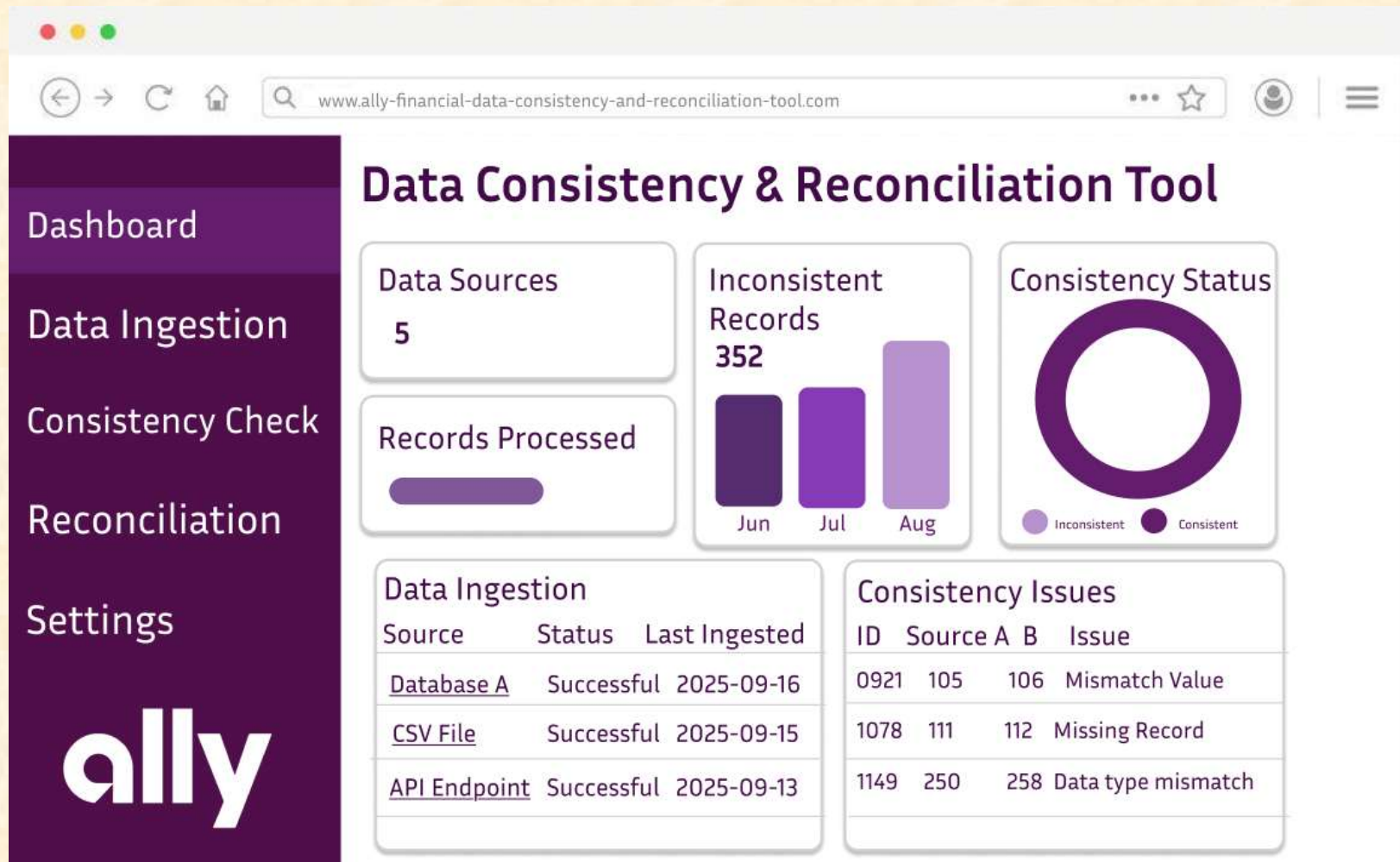
# Project Design Specifications

---

- Dashboard provides insights regarding data analysis and reconciliation.
- Users can upload files in different formats e.g. CSV, XML, excel, etc., for reconciliation.
- Gives users in-depth details regarding rules and discrepancies.
- Users can view log of files and their modification history.

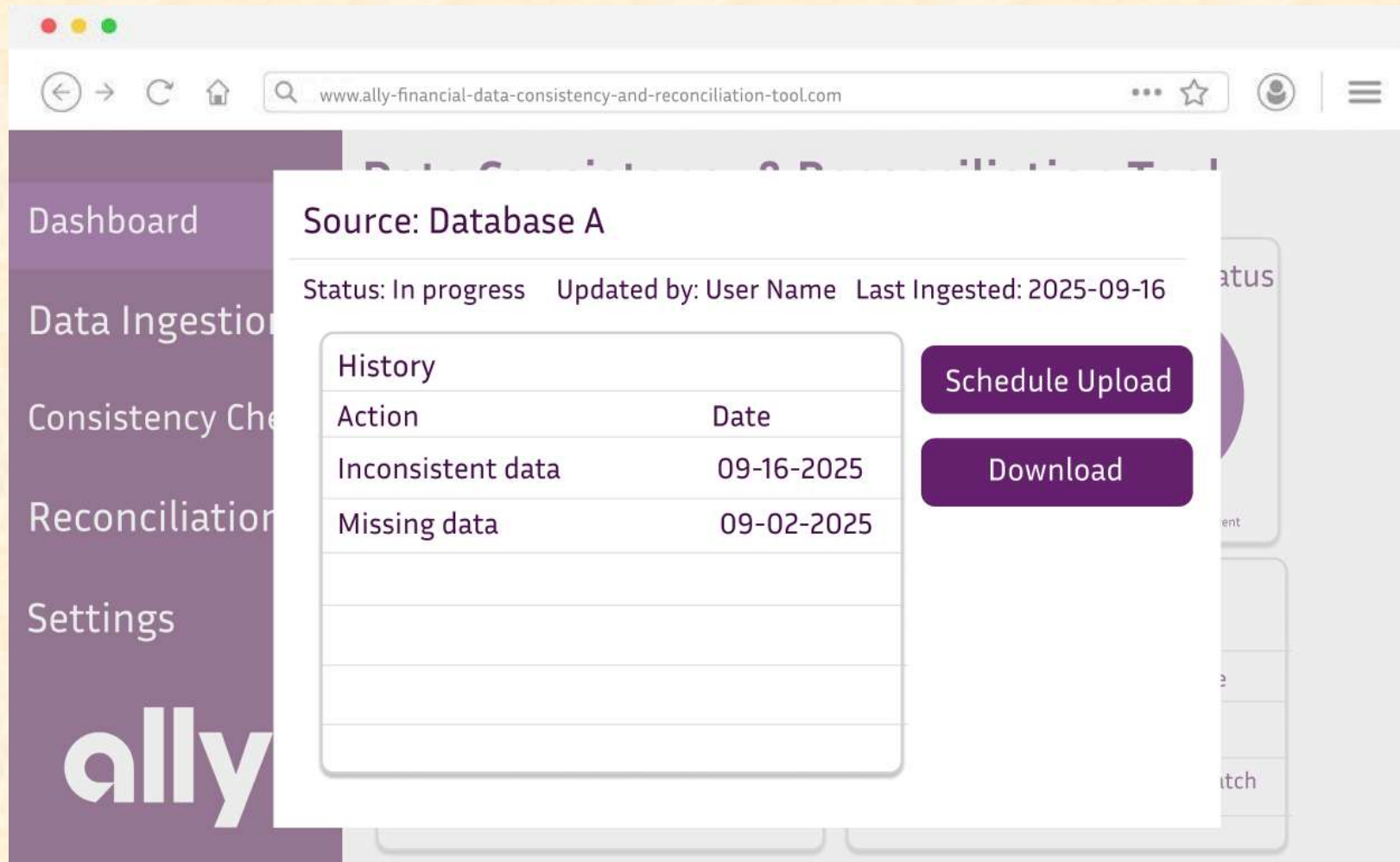


# Screen Mockup: Dashboard





# Screen Mockup: Dashboard Pop-up



# Screen Mockup: Data Ingestion

Dashboard

Data Ingestion

Consistency Check

Reconciliation

Settings

ally

## Data Ingestion

### Source Integration

API Endpoint

CSV File 1

Database A

Database B

### Data Source Upload

Upload a source

No source(s) chosen

### Scheduled Imports

Source	Freq.	Last run	Next run	Status
<a href="#">API Endpoint</a>	Weekly	09-02-2025	09-09-2025	In progress
<a href="#">CSV File</a>	Daily	09-11-2025	09-12-2025	Complete
<a href="#">Database B</a>	Weekly	09-02-2025	09-09-2025	In progress



# Screen Mockup: Ingestion Pop-up

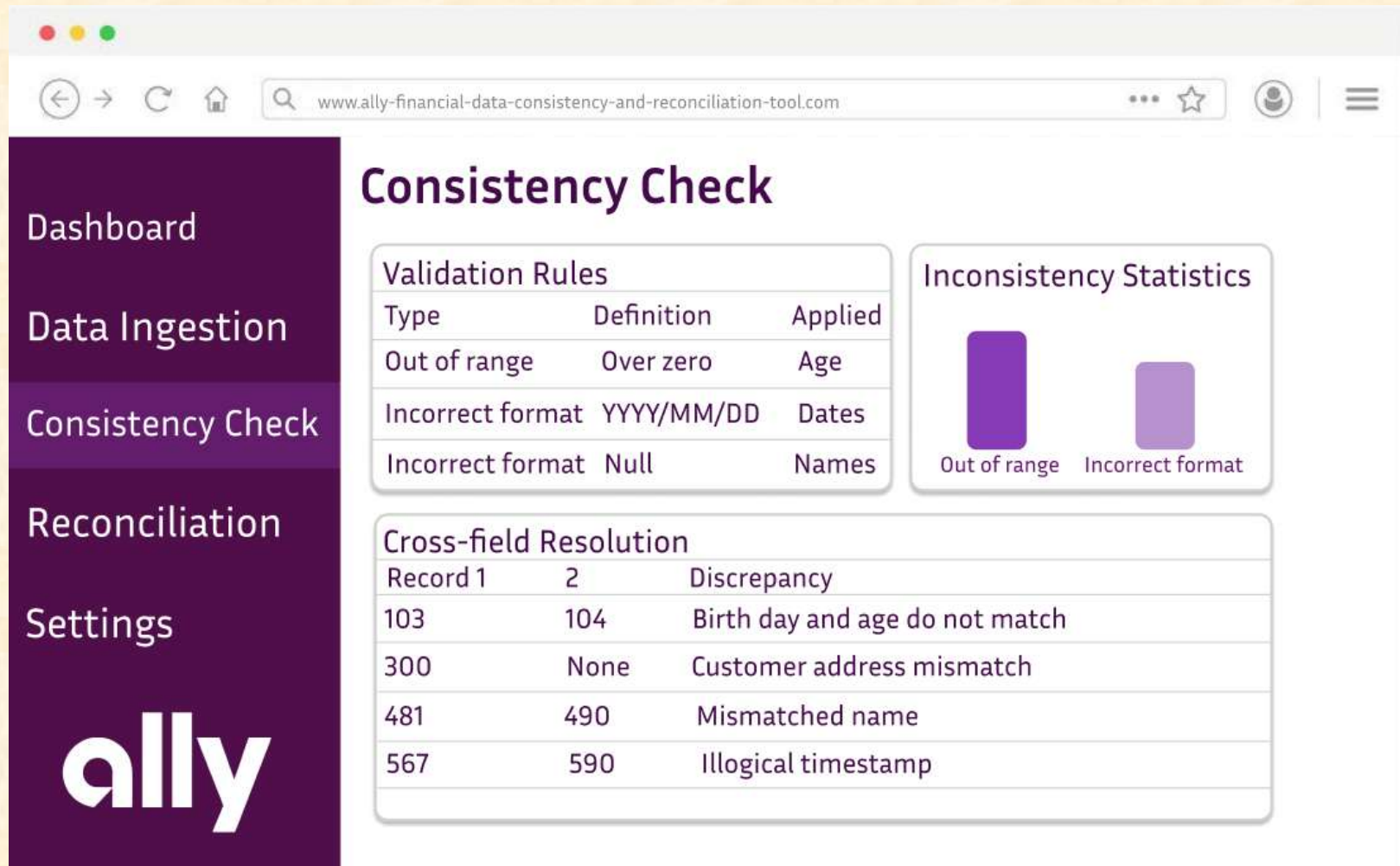
The mockup shows a web browser window with the URL `www.ally-financial-data-consistency-and-reconciliation-tool.com`. A purple sidebar on the left contains the following menu items: Dashboard, Data Ingestion, Consistency Check, Reconciliation, and Settings. The 'ally' logo is at the bottom of the sidebar. A white pop-up window titled 'Upload a source' is centered on the screen. It features a close button (X) in the top right corner. Inside the pop-up, there are two dropdown menus: 'Choose a source type:' and 'Department:'. Below these is a purple 'Upload' button. At the bottom of the browser window, a table displays data for 'Database B'.

Database B	Frequency	Start Date	End Date	Status
Database B	Weekly	09-02-2025	09-09-2025	In progress

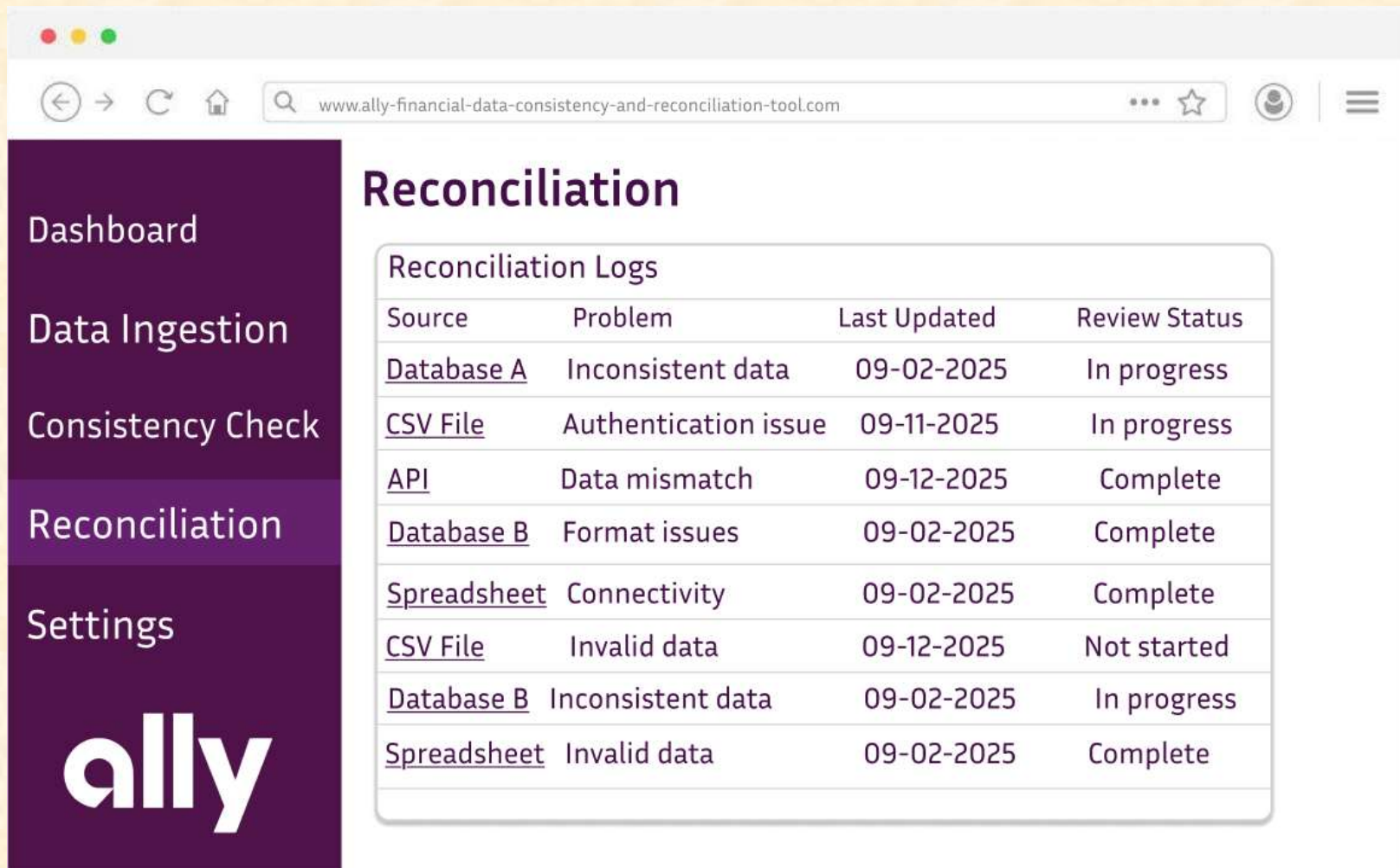




# Screen Mockup: Consistency Check



# Screen Mockup: Data Reconciliation



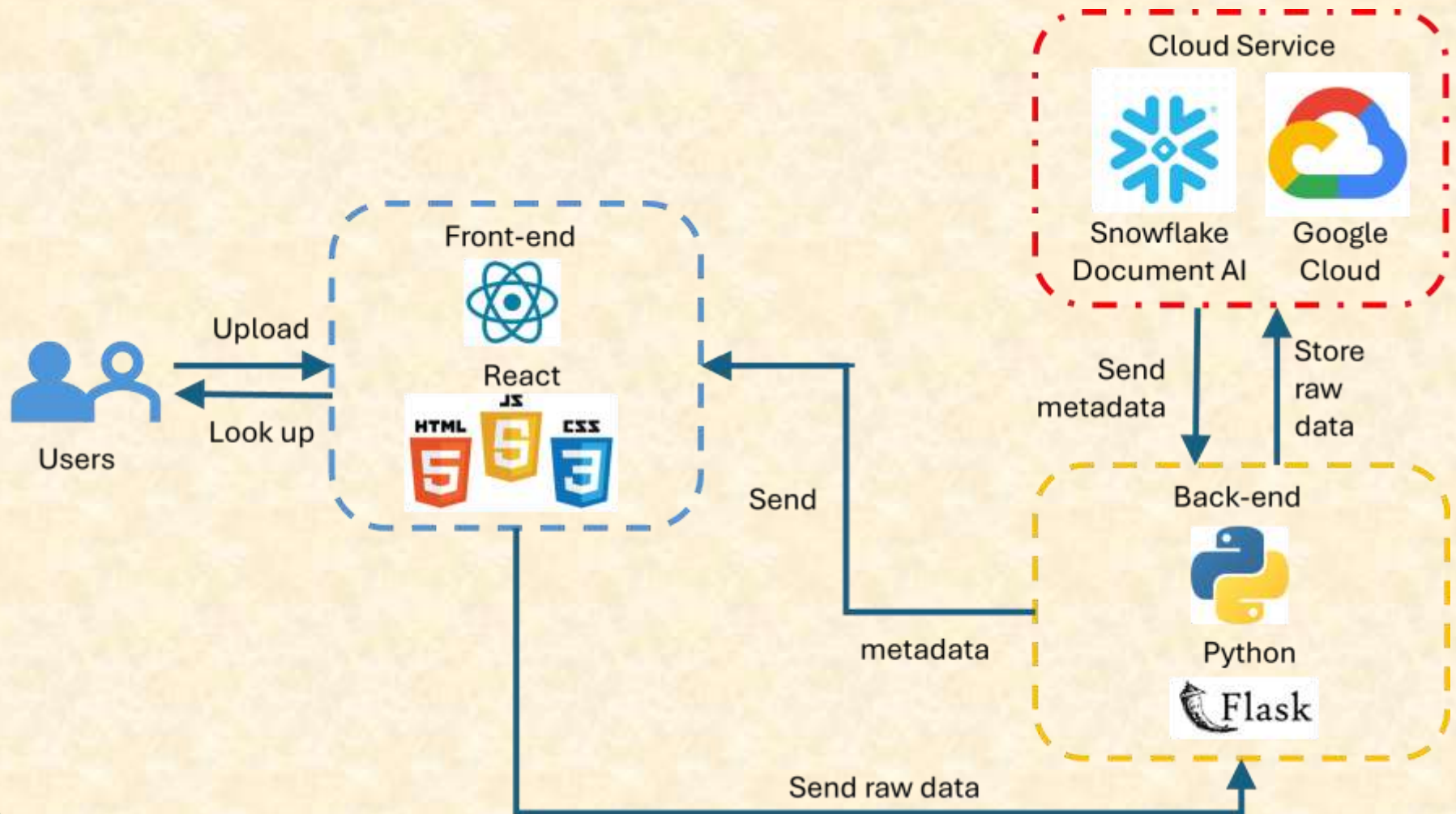
# Project Technical Specifications

---

- React web application
- Python for machine learning tasks
- Flask as backend server
- React frontend, built with Node.js
- Snowflake as data warehouse and processing layer for data
- Google Cloud Storage buckets provided via Snowflake



# Project System Architecture



# Project System Components

- Software Platforms / Technologies
  - Python
  - SQL
  - Flask
  - React.js
  - Node.js
  - Snowflake
  - Google Cloud Platform





# Project Risks

- Data Processing and Storage
  - Upload dataset from user to Snowflake cloud and use SQL to query and send back.
  - Use small test dataset first and split into 3 storage layers (raw, process, metadata) to manage each step within Snowflake no need to send to back-end again.
- Data Format Ingestion
  - Defining the validation rules to deal with different data file format.
  - Allow upload one file at a time and deal with each type separately save it to cloud.
- Data Sources for Testing and Training
  - Using the ML model to look at different content of data sources and creating correlation rules.
  - Familiarize with banking data and efficient methods to utilize ML techniques.
- Data Security
  - Security of data when AI to analyze and process the data.
  - Filter out the confidential information and hide the label of the data and just analyze the consistency of the content type, range, and differences.



# Questions?

---

?

?

?

?

?

?

?

?

?

