

USER DOCUMENTATION

1.Introduction:

The purpose of this tool is to streamline and enhance the methods for interns to explore and search through data related to cars or automotive vehicles. The emphasis is on providing a functional solution with a minimal working example (MVP) that prioritizes functionality over the User Interface/User Experience (UI/UX). The creation of project instances is facilitated by Information Technology (IT) officers using command-line interface (CLI) tools, specifically terraform in this context.

2.User Requirements:

- ***Document Upload and Data Viewing:***
 - Users, identified as IT officers, should have the ability to upload documents in CSV format.
 - The focus is on enabling users to view processed data without a strong emphasis on UI/UX.
 - Any pipeline operation, such as word count, should be seamlessly integrated into the data viewing process.

3. Workflow:

a. User Functionalities:

This section outlines the various functionalities or features from the user's perspective. It implies a user-centric approach to development, emphasizing the importance of understanding and documenting how users interact with the system.

b. Log in:

- Description: Describes the log-in functionality.
- Details: Includes password verification, with a specific mention that unsuccessful password verification leads to an error page. This adds clarity on the expected behavior in case of login failure.

c. Upload Files:

- Description: Describes the file upload functionality.
- Details: Mentions that files are stored in an S3 Bucket. This provides information on the data storage mechanism and where the uploaded files are managed.

d. Add Data Pipeline:

- Description: Describes the process of adding a data pipeline.
- Details: Specifies the use of an external data pipeline and the configuration of Apache Airflow. This indicates the integration of external tools and the specific technology stack being employed.

e. Start Data Pipeline:

- Description: Describes the initiation of a data pipeline.
- Details: Highlights that files in the S3 Bucket are fed into the data pipeline, and the results are stored as metadata in the database. This step-by-step breakdown clarifies the flow of data and the outcomes of the data pipeline process.

f. Search File:

- Description: Describes the file search functionality.
- Details: Includes checking the file in the database and the ability to download the file. This illustrates the database interaction and the retrieval process.