DESIGN DOCUMENTATION

UI Considerations:

- While UI is not the primary focus, a filter search bar is implemented for efficient data retrieval and exploration.
- The search functionality extends to the frontend, allowing users to interact with and visualize the processed data.

Access Management:

- Authentication is handled through Auth0, providing a simple login mechanism.
- Differentiated access levels are not implemented; once a user is registered, they gain access to the entire database, ensuring a streamlined and uniform access control system.

Backend Functionality:

- ***Database Interaction:***
  - The backend interacts with the database, facilitating data storage and retrieval.
  - ***Result Storage:***
    - Processed results are stored in the database in a key-value pair format, ensuring efficient and structured data storage.
- ***S3 Bucket Functionality:***
  - ***File Storage:***
    - The backend is responsible for storing files in the S3 bucket.
    - This includes uploading and managing various types of files within the S3 storage.
  - ***File Retrieval:***
    - Downloading files from the S3 bucket is also a backend-managed process.
    - Users can request files, and the backend facilitates the retrieval and delivery of the requested files.
  - ***AWS Credentials:***

        - Secure AWS credentials are utilized by the backend to access the S3 bucket.

- ***Integration with Apache Airflow:***
    - Apache Airflow is integrated into the system to manage and execute data processing tasks.
    - It receives incoming data, performs operations like word count or key-value operations, and communicates the results back to the backend.

Infrastructure Deployment with Terraform and EKS Cluster:

- ***Terraform Deployment:***
    - The entire system, including the backend, database, and Apache Airflow, is deployed using Terraform.
    - Terraform scripts define the infrastructure as code, allowing for consistent and repeatable deployments.
- ***Elastic Kubernetes Service (EKS) Cluster:***
    - The infrastructure is hosted within an EKS cluster, providing scalability, manageability, and automated Kubernetes orchestration.