





Orquestração de Fluxos de Dados com Apache Airflow

Projeto 3 Arquitetura e Implantação de Real-Time ETL com Airflow, Spark e Kafka



Orquestração de Fluxos de Dados com Apache Airflow

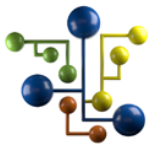
Projeto 3

Arquitetura e Implantação de Real-Time ETL com Airflow, Spark e Kafka

Neste projeto você vai construir e implantar, passo a passo, uma arquitetura de ETL em tempo real, integrando diversas ferramentas. O Apache Airflow será usado para orquestração de workflows, enquanto o Kafka será usado para gerenciar o streaming de dados em tempo real. O PySpark será utilizado para processamento distribuído de dados e o resultado do processamento em tempo real será armazenado em um banco de dados NoSQL.

Todo o pipeline será criado para funcionar de forma automatizada.

Bons estudos.



Equipe DSA

Muito Obrigado!
Continue Trilhando Uma Excelente Jornada de Aprendizagem.