

Abel Mitiku

abel@gettenacious.com · +251919395958

Addis Ababa, Omedla St.

portfolio: <https://abelblue.is-a.dev>

linkedin: <https://www.linkedin.com/in/abel-mitiku-2b95bb215>

github: <https://github.com/abel-blue>

DATA ENGINEER

I am a data engineer/data scientist with technical expertise in a wide range of tools and platforms required for data engineering pipelines, data analysis and data visualization. I have experience with a variety of Amazon tools, including Amazon DynamoDB, Amazon Athena, AWS Glue, Amazon EC2, Amazon Redshift, Amazon S3, and Amazon Web Services (AWS).

KEY SKILLS

Amazon DynamoDB
Amazon Athena, S3
AWS Glue, Redshift
Amazon EC2

ETL, Postgres
OLAP
dimensional modeling
Spark / Airflow / Kafka / Docker

Tableau / Redash
communication

PROFESSIONAL EXPERIENCE

Ubongo
Data Engineer

Mar 2023 - Present

As a data engineer at Ubongo Company, I was responsible for designing and implementing data pipelines that collected and processed data from various sources, including YouTube, Twitter, and Facebook. By leveraging Google APIs and using Amazon tools for scheduling and execution, I helped the company to collect and analyze large volumes of data in a scalable and efficient manner. The goal was to increase viewership and market value and to use data and AI to gain insights into customer behavior, market trends, and business performance, and to use these insights to improve the company's products and services.

Accomplishments:

- Designed and implemented a data pipeline that ingested data from multiple sources including YouTube, Twitter, and Facebook.
 - Leveraged Google APIs to collect data from these sources and ensured the data was cleaned and formatted for analysis.
 - Created automated processes for collecting and processing data, reducing manual effort and increasing efficiency.
 - Worked with Amazon tools such as AWS Glue, AWS Lambda, and Amazon EMR to schedule and execute data pipelines.
 - Built and maintained a scalable infrastructure for storing and processing large volumes of data.
 - Developed data models and implemented best practices for data management and quality assurance.
 - Conducted performance analysis to optimize data pipelines and improve overall system efficiency.
-

AI Guild
Data Engineer/AI Practitioner

Dec 2022 - Sep 2023

We aim to promote ethical and responsible AI practices and help organizations harness the potential of AI to solve real-world problems.

- promoting ethical and responsible AI practices and helping individuals and organizations develop the skills and knowledge they need to succeed in the rapidly-evolving field of AI.
-

PROJECTS

Speech data collection using kafka cluster

- By combining messaging, storage, and stream processing, the data pipeline allow collection, storage and analysis of real-time audio datasets.
- Used techstacks: AWS S3 Bucket, aws athena, spark, kafka, ELT pipeline, Airflow, docker

Data warehouse for pNEUMA | open-traffic datasets

- Built a scalable data warehouse tech-stack with ELT pipeline composed of MYSQL Database, dbt transformation, Airflow orchestration and Redash visualization.
- Used tech-stacks: dbt, ELT pipeline, Airflow, Redash.

USGS LIDAR Data warehouse

- produce an easy-to-use, reliable and well-designed Python module that domain experts and data scientists can use to fetch, visualize, and transform publicly available satellite and LIDAR data. In particular, our code should interface with USGS 3DEP and fetch data using their API.
- Used: LASzip, PDAL, CloudCompare, USGS 3DEP API, GIS software packages such as QGIS or ArcGIS.

Speech Recognition

- Built end-to-end speech recognition model that transcribes Amharic speech to text
- Used ML frameworks like TensorFlow and PyTorch and NLP
- Used techniques like Fourier transforms, short-time Fourier transforms, Mel-frequency cepstral coefficients (MFCCs).

EDUCATION

Addis Ababa University 2016-2022

Computer Engineering

Ten academy 2021-2022

Data Science
