OCHIENG HERMAN OTIENO

Phone: +254 743 851 498 | **Email:** ochiengherman32@gmail.com | **Address:** Nairobi, Kenya
LinkedIn: linkedin.com/in/ochieng-herman-9605632a0
GitHub: github.com/ochiengherman36

PROFESSIONAL SUMMARY

Backend Data Engineer with 4+ years of hands-on experience building ETL/ELT pipelines, backend APIs and reporting infrastructure for data-driven operational decisions. Experienced across the full data lifecycle, from raw production data ingestion through transformation, modelling and delivery into clean, documented and query-ready datasets and dashboards. Built AI-powered applications using RAG, LangChain, LangGraph and multi-agent orchestration, including hybrid document retrieval and real-time recommendation systems, alongside backend development using FastAPI, Django, Laravel and Express.js. Seeking to apply expertise in data engineering and backend development to Data Engineer, Backend Engineer or AI Engineer roles requiring scalable pipelines, reliable APIs and intelligent data applications.

SKILLS

Data Engineering, Backend Engineering, Pipeline Development, Data Modelling, ETL Development, API Development, SQL Development, Database Management, Data Integration, Data Streaming, Data Warehousing, Data Governance, Change Capture, Batch Processing, Stream Processing, Query Optimisation, Retrieval Systems, Vector Search, Machine Learning, Model Explainability, Technical Documentation, Reporting Automation, Dashboard Development, Cloud Deployment, Server Management, Distributed Computing, Multi-Agent Systems, Recommendation Systems, Predictive Modelling, Data Analytics.

EXPERIENCE

Backend Data Engineer
Data Science East Africa | Nairobi, Kenya | April 2023 – Present
Roles and Responsibilities:

- Built and maintained ETL/ELT pipelines that moved production database data into analytics-ready stores supporting reporting and operational decision-making.
- Developed SQL queries and reporting workflows that converted operational data into actionable outputs for time-sensitive business decisions.
- Automated scheduled reporting pipelines to improve consistency and accessibility of recurring operational reports.
- Created Power BI and Apache Superset dashboards that provided stakeholders with self-service visibility into business and operational data.
- Documented data logic, dependencies and reporting workflows to improve maintainability and knowledge transfer across technical and non-technical teams.
- Coordinated reporting documentation across distributed stakeholders, maintaining clear records of pipeline logic, queries and dashboard dependencies.
- Supported cross-functional data initiatives by aligning technical reporting requirements with stakeholder needs across multiple client projects.

Backend Data Engineer
Upwork | Nairobi, Kenya | Remote | September 2022 – Present
Roles and Responsibilities:

- Developed backend services for international clients using Laravel, Django and FastAPI across remote-first development environments.
- Engineered backend data models and exposed structured datasets through REST APIs for integration with application and reporting workflows.
- Built frontend applications using React and TypeScript, integrating user-facing interfaces with backend API services.
- Provisioned and managed Linux servers, including database configuration, setup and remote access required for deployment.
- Produced technical documentation covering data analysis and data engineering workflows to support distributed technical teams.
- Delivered backend engineering services remotely for Owlswift Commerce, supporting an international e-commerce business selling home-based products.

- Integrated backend, database and deployment components across virtual development workflows to support reliable delivery for geographically distributed clients.

PROJECTS

Enterprise Knowledge Assistant — RAG + Multi-Agent AI

- ❖ Built a hybrid RAG system for grounded question answering across enterprise PDF documents.
- ❖ Implemented semantic retrieval through ChromaDB alongside keyword search through Elasticsearch to improve document discovery.
- ❖ Combined retrieval results using reciprocal rank fusion, producing stronger retrieval performance than either search approach independently.

Retail Recommendation Platform

- ❖ Built an end-to-end retail recommendation pipeline processing streaming order events through Kafka producers and consumers.
- ❖ Implemented Parquet-based data lake storage for streamed retail transaction and interaction data.
- ❖ Developed PySpark ETL workflows that cleaned and encoded customer interaction data for recommendation modelling.

End-to-End E-Commerce Data Pipeline

- ❖ Built an end-to-end data pipeline covering raw data ingestion, change-data capture, transformation and analytical modelling.
- ❖ Implemented Kafka-based streaming workflows for processing incoming e-commerce data.
- ❖ Integrated Debezium to capture database changes and support continuous data movement.

Maize Lethal Necrosis Disease Detection (ML + LLM)

- ❖ Built an imbalance-aware stacked ensemble for detecting maize disease from leaf-symptom data.
- ❖ Applied scikit-learn to develop and evaluate machine learning workflows.
- ❖ Integrated XGBoost, LightGBM and TabNet into the predictive modelling pipeline.

EDUCATION

- ❖ **Bachelor's Degree, Computer Science |** Laikipia University
- ❖ **Data Engineering |** LuxDevHQ Academy

TOOLS

Python, SQL, PL/SQL, PHP, TypeScript, Java, dbt, Apache Airflow, Apache Spark, PySpark, Apache Kafka, Debezium, FastAPI, Django, Laravel, Express.js, Spring Boot, REST APIs, React, PostgreSQL, SQL Server, MongoDB, Cassandra, AWS, Google Cloud Platform, Docker, Linux, Power BI, Apache Superset, Elasticsearch, ChromaDB, LangChain, LangGraph, RAG, PyTorch, scikit-learn, XGBoost, LightGBM, TabNet, SHAP, Claude, Git.

LANGUAGES

- ❖ English - Fluent
- ❖ Swahili - Native

REFEREES

Available Upon Request.