

XAVIER FRANKLINE ODHIAMBO

xavierfrank4@gmail.com | xodhiambo.com | github.com/theplsalmist | +254786409255 | Remote

Summary

- Senior software engineer with 7 years of Python experience building production web applications across SaaS, civic-tech, and distributed data platforms. Strong on API engineering (Django REST Framework, FastAPI), database design and optimisation (PostgreSQL, ClickHouse, Elasticsearch), cloud infrastructure (AWS Solutions Architect certified), and operating distributed systems through CI/CD, observability, and incident response. Recent work includes production AI engineering — RAG pipelines, LLM orchestration infrastructure, and self-hosted model deployment. Open-source contributor to Django and CKAN.

Skills

Languages: Python, TypeScript, SQL, Bash

Backend & APIs: Django, Django REST Framework, FastAPI, Flask, Celery, RESTful API design, validation, authentication, API security

Databases: PostgreSQL (Django ORM, migrations, indexing, query optimisation, schema design), ClickHouse, Elasticsearch, Redis

Cloud & Infrastructure: AWS (EC2, RDS, ECS, Lambda, S3 — Solutions Architect certified), Docker, Kubernetes, Ansible, Terraform

CI/CD & Tooling: GitHub Actions, GitLab CI/CD, Git workflows, code review

Observability: OpenTelemetry, Grafana, Loki, Tempo, Prometheus, Sentry

Data Engineering: ETL pipelines (dbt, Stitch), distributed worker pipelines (RabbitMQ, Celery), data validation and normalisation at scale

AI Engineering & Tools: RAG pipelines (LlamaIndex, pgvector), LLM orchestration (OpenRouter multi-provider routing), self-hosted model deployment (Ollama, OpenWebUI), prompt engineering and templating, AI coding tools (Claude Code, Cursor)

Certifications

AWS Certified Solutions Architect – Amazon Web Services

January 2025

AWS Certified Cloud Practitioner – Amazon Web Services

January 2022

Open Source Contributions

Django: Bug fixes and documentation contributions to the framework.

CKAN: Extension and plugin contributions to the leading open-source platform powering government and institutional open-data portals worldwide.

Experience

Senior Software Engineer, Search Atlas — Remote

August 2025 – June 2026

Backend engineer on an SEO analytics platform integrating Google Search Console, Google Analytics, Google Ads, and Meta Ads for advertising agencies and in-house marketing teams.

- Designed and built REST API services in Django REST Framework with strong input validation, structured error handling, authentication, and rate limiting — backing both internal services and a customer-facing analytics product.
- Wrote and shipped PostgreSQL and ClickHouse migrations, optimised slow queries through indexing and schema refactors, and worked on partitioning strategies for high-volume analytics tables.
- Instrumented services with OpenTelemetry across Django and Celery; resolved production incidents by correlating Sentry exceptions, Loki logs, and Grafana metrics.
- Built and maintained the platform's LLM orchestration layer (multi-provider routing, automatic failover, cost-aware model selection) used by multiple internal services for automated analysis and content generation — see Selected Projects.

Senior Software Engineer, Code for Africa — Remote

August 2022 – August 2025

Backend and infrastructure engineer building systems for civic-tech platforms used by researchers, journalists, and policy teams across Africa.

- Built and maintained Django REST Framework, FastAPI, and Flask APIs across multiple platforms — including distributed news indexing, IoT sensor ingestion (Sensors.Africa), and census/demographic data platforms (Wazimap-NG).
- Designed and deployed multi-node Elasticsearch clusters using Ansible automation, and implemented end-to-end observability with Grafana, Prometheus, Loki, and Alertmanager — investigating and resolving production incidents by correlating Sentry exceptions, logs, and metrics.
- Drove adoption of CI/CD pipelines across engineering teams using GitHub Actions and GitLab CI/CD, with automated testing, containerised image publishing, and technical documentation as standard practice.
- Led development of internal AI tooling, including a multi-stage RAG-based content classifier and a self-hosted LLM platform serving non-engineering teams — see Selected Projects.

Software Engineer, QED Solutions Ltd. — Nairobi

April 2021 – July 2022

Backend engineer on Tendersure Africa, a SaaS B2B procurement platform used across multiple African countries.

- Built and maintained the platform with Django (backend) and Vue.js (frontend), integrating localised payment gateways across multiple African markets — handling signed callback verification, idempotent retries, and reconciliation.
- Migrated legacy applications from VM-based workloads to Docker containers on Amazon EC2, decoupling deployment from infrastructure.

Software Engineer, Nakala Analytics Ltd. — Nairobi

August 2020 – March 2021

- Built data analytics applications in Python and Flask with REST APIs over PostgreSQL, integrating data from third-party APIs for downstream visualisation.
- Built ETL pipelines integrating PostgreSQL, dbt for transformation, and Stitch for source replication.
- Architected and deployed JupyterHub using Docker and Docker Swarm for the data analytics team.

Software Engineer, JobsearchKE — Nairobi

September 2019 – August 2020

- Built the company's core platform jobsearchke.com using Django, PostgreSQL, and JavaScript.
- Architected and deployed application infrastructure on DigitalOcean using Docker, Nginx, and DigitalOcean Spaces.
- Built an automated campaign and marketing platform integrating Mailchimp and SendGrid APIs.

Selected Projects

Trustlist (Python, LlamaIndex, pgvector). Led development of an internal ad-quality and domain-trust classifier — a gated multi-stage RAG pipeline that scored advertisement URLs through cheap signals first (domain popularity, infrastructure trust checks, link-graph authority), then sent surviving URLs through an LLM content-quality stage that embedded the scraped page and queried for trust-indicator markers through templated prompts. The cheap-signals-first gating reserved LLM cost and latency for URLs already worth investigating, not spent uniformly across the input.

Mediacloud Story-Indexer (Python, FastAPI, RabbitMQ, Elasticsearch). Contributed to an open-source distributed news ingestion and indexing pipeline tracking thousands of news sources globally and indexing millions of stories for researchers — FastAPI services for orchestration, Elasticsearch query layers for downstream tooling, and operational work on a multi-node Elasticsearch cluster deployed via Ansible automation.

LLM Orchestration Layer at Search Atlas (Python, OpenRouter). Built and maintained the orchestration layer that multiple internal services depended on for automated analysis and content generation — multi-provider routing, automatic failover, cost-aware model selection, and structured logging the team could use to reason about model behaviour over time.

Self-hosted LLM Platform at Code for Africa (Ollama, OpenWebUI). Designed, deployed, and operated a self-hosted LLM platform that gave non-engineering teams safe, private access to AI tools for sensitive workloads — with access controls, model selection guidance for non-technical users, and ongoing adoption support.

Education

Jomo Kenyatta University of Agriculture and Technology — BSc Telecommunications Engineering November 2018