

Paul Mutemi
Cloud Engineer | Software Engineer
Phone: +254718769551 **Email:** laupmutemi@gmail.com
LinkedIn: <https://www.linkedin.com/in/paul-mutemi/> **Github:** <https://github.com/mutemip>

PROFESSIONAL SUMMARY

Dynamic Software Engineer with expertise in backend, and cloud engineering . Proven ability to deliver scalable software solutions and optimizing system performance. Experienced in building high-throughput production applications serving millions of daily requests. Excels in agile environments, contributing to project success through effective communication and teamwork.

KEY TECHNICAL SKILLS

Backend: Python, Django, Django REST Framework, FastAPI, Flask .
Databases: PostgreSQL, MySQL, MongoDB .
AI/LLM Integration: OpenAI, Claude, Gemini APIs, RAG pipelines, Agentic AI
Cloud and DevOps: AWS, Azure, GCP, Docker, Kubernetes, GitLab CI/CD, Terraform.
Frontend: React, Next.js, React Js, HTML, CSS, Tailwind CSS .
Tools and Libraries: Git, RabbitMQ, Celery, Kafka, Redis

WORK EXPERIENCE

New Age - (Remote)

Backend Software Engineer

June 2024 – Present

- Leads a team of backend engineers to design and ship high-throughput APIs, delivering core business logic within a distributed microservices that serves millions of requests per day.
- Design, build, and deploy a performant, reliable, scalable, and fault-tolerant cloud native microservice based web service on the cloud within a specified budget.
- Develops scalable APIs for a Retrieval-Augmented Generation (RAG) system, integrating vector embedding models to enhance user query responses with information from a large-scale knowledge base.
- Design, implement, test, package, deploy, and monitor cloud applications using Virtual Machines (VMs), and Containers cloud computing services.

PawaIT Solutions Ltd - Nairobi, Kenya

Software Engineer

Feb 2024 – May 2024

- Implemented an automated data processing pipeline in Python for pre-processing large volumes of audio data, which streamlined the AI transcription using whisper large model and ground-truthing workflow by 45% .
- Designed and implemented backend for an AI customer insight platform, integrating the OpenAI Whisper model and Langchain to orchestrate an AI pipeline for automated audio transcription and sentiment analysis.
- Developed a highly scalable, distributed task queue system using Celery and RabbitMQ, and integrated a Redis-based cache that slashed processing latency by 40% while handling over thousands of concurrent ffmpeg audio processing background tasks per minute.
- Designed, implemented, tested, and debug applications using interactive, batch, and stream processing frameworks and compare their suitability to different problem domains.

United States International University - Africa - Nairobi, Kenya

IT support Specialist

April 2023 – Feb 2024

- Led cross-functional team initiatives that revolutionized systems operations, achieving 25% improvement in overall efficiency.
- Created automated reporting scripts that consolidated system metrics from multiple sources, reducing report generation time by 60%
- Configured user accounts management on Active Directory to enforce least privilege on critical infrastructure, enabling network resource sharing among various organization departments.

Power Learn Project - Nairobi, Kenya

Lead Python Instructor

May 2022 – April 2023

- Led pan-African virtual mentorship program across five countries(Kenya, South Africa, Nigeria, Zambia and Tanzania), implementing structured learning pathways that elevated program graduation rates by 20%.
- Organized quarterly hackathon events engaging 500+ participants, establishing competitive challenges and collaborative projects that accelerated practical skill development.
- Designed and developed a comprehensive software engineering curriculum for a coding bootcamp, covering full-stack development, best practices in version control, agile methodologies, and real-world project implementation.
- Conducted code reviews and mentored students on best practices for clean, maintainable code; facilitated training sessions on contributing to open-source projects, including Git workflows, issue tracking, and pull request etiquette

Volunteering

Black Python Devs Open-Source Community (Remote) – Florida, USA

Open-Source Community Mentor

July 2024 – November 2024

- Led mentorship initiatives for 10+ developers, establishing structured learning paths for open source contribution and GitHub workflow best practices
- Planned hands-on workshops on Git fundamentals, pull request workflows, and collaborative development practices, resulting in 100+ successful first-time contributions.
- Guided developers through their first pull requests, achieving an 80% successful merge rate and fostering long-term community engagement.

EDUCATION - Remote

Carnegie Mellon University Africa

- **Master of Science in Information Technology**

Karatina University – Kenya

- Bachelor of Science in Computer Science.

Professional Certification

- Cloud Developer - Udacity
- Cyber Security Analyst - CyberShujaa Bootcamp.
- Meta Backend Professional Certificate - Coursera/Meta.

Projects

1. [Grammar AI checker and corrector - IOS Application](#) - (An App on App Store)

- A productivity application to elevate user English grammar which acts as your sentence corrector, plagiarism checker, spell check, document editor and punctuation checker. The Application has approximately over 532,000 users.
 - Led design and implementation of backend logic by integrating AI models(OpenAI models, ClaudeAI models and Google Gemini) for text and document processing endpoints.
 - Implemented grammar and spell check functionality, some of key features within the application.
 - Deployed APIs on GCP (Main) and Azure (backup).

Stack and Tools: Python, Django Rest Framework, GCP, Azure, PostgreSQL, AI models, Celery, Redis, RabbitMQ

2. [Currency Converter - IOS Application](#) - (An App on App Store)

- Currency Converter Calculator equips users with live exchange rates, both online and offline mode, making it perfect for travelers. Solves the problem of frantic internet searches or puzzling over unfamiliar currencies.
 - Developed an Object Character Recognition(OCR) pipeline for image processing using Azure vision for computer vision.
 - Implemented an optimized image processing pipeline to streamline a section OCR workflow within the application.
 - Integrated live currency rates APIs for updated currency rates within the application.
 - Deployed APIs backend logic on GCP.

Stack and Tools: Python, Django Rest Framework, GCP, Azure, PostgreSQL, AI models, OpenCV, Computer Vision, Celery, Redis, RabbitMQ

3. [Obi Avartar](#)

- Ms. Proactive is Desktop AI assistant (PyQt5) that answers from a PDF knowledge base using RAG (ChromaDB + local embeddings), Ollama for the LLM, Whisper for speech-to-text, neural TTS, and Wav2Lip for lip-synced avatar video, designed to run fully on-device without paid API subscriptions.
 - Architected and built a local-first conversational desktop app that chains retrieval-augmented generation (PDF chunking, vector store, LLM context) so answers stay grounded in user documents instead of generic model knowledge.
 - Implemented a RAG pipeline with ChromaDB and sentence-transformers embeddings: ingest PDFs (via PyMuPDF), chunk and index text, and retrieve top-k passages to constrain Ollama responses.
 - Integrated a local LLM via Ollama (HTTP client, configurable models) with a streaming-oriented client layer, keeping inference on-machine and API-key-free for core chat.

Python · PyQt5 · RAG · ChromaDB · sentence-transformers · Ollama · Whisper · edge-tts · Wav2Lip · OpenCV · PyMuPDF · FFmpeg

4. [MyPowerly](#) (Ongoing Project)

- A complete e-signature workflow and a sales tax returns application for residents in various states in the US.
 - Implemented tax return api endpoints for handling sales tax return for individual, s-corporation, c-corporation, partnerships, fiduciaries and exempt organizations in various US states.
 - Integrated e-signature processing functionality for files from local machine, Google storage and Microsoft 365.

Stack and Tools: Python, FastAPI, Celery, Redis, RabbitMQ, AWS, PostgreSQL