

STEPHEN KAMAU

AI Specialist & Software Developer | Python & Systems Architecture

kamstefin@gmail.com | +254 115 560 414 | Portfolio: fskamau.vercel.app

LinkedIn: linkedin.com/in/fskamau | GitHub: github.com/fskamau | Nairobi, Kenya

PROFESSIONAL SUMMARY

Versatile AI Specialist and Software Developer with 5+ years of practical engineering experience spanning production software development, LLM benchmark evaluation, and multimodal AI training pipelines. Proficient in architecting modular backend services in Python, containerizing workflows via Docker, and building deterministic evaluation harnesses for autonomous agents. Expert in auditing code quality, analyzing system failure modes, and standardizing benchmark datasets that align generative models with rigorous production engineering standards.

CORE TECHNICAL COMPETENCIES

- **AI Systems & Evaluation:** LLM code and reasoning evaluation, Computer-Use Agent (CUA) benchmarking, Side-by-Side (SxS) visual QA, function-calling verification, rubric design
- **Languages & Frameworks:** Python (5+ yrs), JavaScript/TypeScript, FastAPI, Django, REST APIs, JSON/YAML schema validation, SQL, Bash
- **Computer Vision & Annotation:** Multi-frame video object tracking, polygon segmentation, temporal event boundary tagging, CVAT, OpenCV (CV2)
- **Architecture & Quality Engineering:** Modular system design, SOLID principles, root-cause debugging, AST code analysis, pytest, Bandit, Flake8, CI/CD
- **Databases & DevOps:** PostgreSQL, MySQL, SQLite, Redis, Docker (containerization), Linux/Unix CLI, Git/GitHub collaborative workflows

PROFESSIONAL EXPERIENCE

Turing — Senior AI Specialist & Code Quality Evaluator

May 2025–Present

Contract | Remote

- Review, analyze, and rewrite AI-generated codebases across Python, TypeScript, and SQL to meet production engineering standards.
- Perform deep root-cause debugging on failing edge cases, refactoring brittle scripts into hardened, scalable, and modular reference implementations.
- Evaluate multi-step action trajectories of autonomous Computer-Use Agents (CUA) across terminal shells, databases, and desktop tools.
- Formulate objective scoring rubrics, benchmark datasets, and technical annotations to fine-tune frontier coding models.
- Verify function-calling parameters, intermediate application states, and tool-use executions against rigid technical criteria.

- Author structured technical documentation outlining failure modes, architectural anti-patterns, and recovery strategies.

CloudFactory — Software Developer & Data Pipeline Engineer

January 2020–April 2025

Remote | Full-time

- Built, deployed, and maintained production-grade Python web applications and automated data-processing pipelines.
- Led technical code reviews, enforcing clean architecture patterns, PEP8 compliance, strict type hinting, and unit test coverage.
- Containerized local application workflows and database instances using Docker, ensuring environment parity across distributed teams.
- Optimized relational database queries, schema migrations, and indexing strategies across PostgreSQL and MySQL databases.
- Engineered robust schema validation routines using Pydantic and pytest to eliminate unhandled data ingestion errors.
- Mentored developers on defensive programming, test-driven development, and collaborative Git/GitHub workflows.

EDUCATION

Bachelor of Science in Computer Science

September 2021–May 2025

Meru University of Science and Technology

CERTIFICATIONS & TECHNICAL CREDENTIALS

Advanced Software Architecture & Engineering Program

June 2025

Professional Engineering Certification

- Professional Python Software Architecture & Design Patterns (Issued July 2025)
- Secure Coding & Software Vulnerability Remediation (OWASP) (Issued August 2025)
- LLM Evaluation & Benchmarking Methodologies (Issued September 2025)
- Docker Certified Associate — Containerization & Environment Isolation