

FELIX KWASI SARPONG

Lubbock, TX (Open to Remote/Hybrid/Relocation) | felixsarpong25@gmail.com | linkedin.com/in/felixsarpong | github.com/felixsarpong | felixsarpong.dev

SUMMARY

Backend-first Software Engineer (5+ years) building Java/Spring Boot and Python/FastAPI systems on AWS for fintech/payments. Delivered event-driven microservices (Kafka/JMS) at scale (1M+ tx/day; 20M+ tx/month) with strong reliability (MTTR < 35%) and performance (p95 < 140ms). Builds LLM products (RAG, tool-calling) with governance and grounding/citations across OpenAI/Anthropic and AWS Bedrock.

SKILLS

Java, Python | Spring Boot/Security (OAuth2/JWT), FastAPI, REST/OpenAPI, GraphQL | Kafka, JMS
PostgreSQL, Oracle, MySQL, Redis | Docker, Terraform, GitHub Actions | AWS (ECS/Fargate, Lambda, SQS/SNS, RDS, S3, IAM, CloudWatch, Bedrock)
Observability: Prometheus/Grafana, Micrometer | Testing: JUnit, Pytest | Workflow: Apache Airflow
AI/Agents: LangGraph, CrewAI, LangChain, LlamaIndex, MCP, Neo4j, Faiss, Pinecone | LLMs: OpenAI, Anthropic, Ollama

EXPERIENCE

Machine Learning & Generative AI Engineer - Tata Consultancy Services (TCS) | Dallas, USA (Remote) | Jan 2024 - Dec 2024

- Built and deployed LLM-powered assistants for enterprise workflows (case triage, knowledge search, summarization) using RAG with citation/grounding to reduce hallucinations.
- Designed inference pipelines (prompting + retrieval + tool-calling) with latency/throughput optimizations (caching, batching, streaming responses).
- Implemented evaluation harnesses (offline + A/B-style experimentation) measuring answer quality, retrieval hit-rate, and safety regressions; iterated prompts/retrieval based on metrics.
- Shipped API-first microservices (FastAPI/Node.js) integrating OpenAI/Anthropic, AWS Bedrock, and open-source models; added guardrails (PII redaction, policy checks) and operated production with monitoring, runbooks, retries/timeouts/fallbacks; orchestrated workflows with Airflow.

Software Engineer - EFT Corporation (Payments Platform) | Distributed | Aug 2021 - Dec 2023

- Owned backend services for a prepaid card platform processing 20M+ transactions/month; reduced processing latency 15-25% through Redis caching and SQL/JDBC tuning; implemented OAuth2/JWT RBAC and PCI-aligned controls; supported Visa/Mastercard/Verve readiness and ISO 8583 cutovers.

Software Engineer (Integration) - Republic Bank (Core Banking) | Accra | Oct 2020 - Jul 2021

- Built Temenos T24 integration APIs using SOAP (WSDL)/XML supporting ~7M transactions/month; improved reliability via strict validation, operational logging, and release stabilization during cutovers.

Backend Developer (Contract) - Passmall (E-commerce) | Remote | Dec 2018 - Aug 2021

- Developed FastAPI services for catalog/onboarding/checkout; improved deployment reliability via containerization and CI checks; implemented concurrency-aware real-time integrations with back-pressure patterns.

SELECTED PROJECTS

Sentinel (PyPI: sentinel-core-fks) - Human-in-the-Loop Control Plane for Tool-Using Agents | FastAPI, GraphQL, Postgres, Neo4j, MCP

- Policy-gated tool execution: agents proposeToolCall(tool,args) and policy returns ALLOW / BLOCK / APPROVAL_REQUIRED; approval workflow with audit-grade decision traces.
- Multi-orchestrator + multi-runtime support (langgraph / crewai / autogen; Ollama/OpenAI/Anthropic) with tool discovery/contract sync via syncMcpTools.

Canon - Citation-first academic RAG platform (AWS deployed) | FastAPI, LangGraph, Postgres, FAISS

- Built course-aware, citation-first RAG with deterministic routing; deployed to ECS Fargate + RDS with CI/CD and optional local inference via Ollama.

EDUCATION

M.S., Computer Science - Texas Tech University | 2024-2025

B.S., Computer Science - KNUST | 2016-2020