

Mohamad Kansa

Python, Data & AI Engineer · LLM Applications · Data Engineering · AI Automation

[kansa1245@gmail.com](#) [github.com/MohamadKansa](#) [linkedin.com/in/mohamadkansa](#) London, UK

PROFILE

Python, Data & AI Engineer with commercial experience at Cognizant across RAG-based enterprise search, LLM-backed automation workflows and production data pipelines for UK government clients. Hands-on building AI agents over enterprise data using MCP-style approved tool layers, RBAC-controlled access, FastAPI orchestration, PostgreSQL/Redis-backed traceability, source-grounded retrieval and structured outputs. Delivered a **40% reduction in data latency** across **50M-record** Databricks pipelines and built privacy-first local LLM systems. Strong across Python, SQL, PySpark, LangChain/LangGraph, RAG, vector databases, Azure, Docker and CI/CD.

TECHNICAL SKILLS

Primary Stack	PythonSQLFastAPIREST APILangChainLangGraphRAG Pipelinespytest	GenAI / LLMs	MCP-style ToolsAI AgentsTool CallingRBACFAISSChromaDBEmbeddingsChunkingVector SearchPrompt EngineeringLocal LLMs
AI / ML	PyTorchpandasNumPyscikit-learnXGBoostLightGBMMLflowAgent EvaluationTime-Series Modelling	Data / Cloud	PySparkDatabricksAzure Data FactoryPostgreSQLRedisKeycloak/SSOETL/ELTData PipelinesData Quality
MLOps / DevOps	DockerGitHub ActionsCI/CDGitBashLinuxLogging/TracingModel EvaluationContainerisation	Frontend / Tools	ReactStreamlitWebRTC VADPyAudioffmpegMarkdown Reporting

EXPERIENCE

Cognizant Graduate, Artificial Intelligence & Analytics London, UK · Hybrid	May 2025 – Present
- Developed RAG-based enterprise search and retrieval Q&A prototypes, designing metadata-aware chunking, embeddings, vector search, source grounding, access controls and retrieval evaluation across enterprise knowledge bases. - Designed and built an end-to-end MCP-style AI assistant prototype over structured enterprise data using FastAPI orchestration, approved tool calls, RBAC checks, PostgreSQL, Redis, structured outputs and traceability; independently hardening it toward production readiness with controlled data access, safer tool use and human review. - Maintained and improved production ADF-to-Databricks pipelines for a UK public-sector valuation client, adding retry logic and observability to deliver a 40% reduction in data latency across 50M-record datasets (PySpark, SQL). - Worked on tax-data pipelines and transformation workflows for a UK public-sector client, adding data quality checks, validation and output tracing using Azure, SQL, Python, PySpark and Databricks for source-to-reporting assurance.	
Independent ML & AI Engineering Self-Directed Project Work London, UK · Remote	Sep 2024 – Apr 2025
- Built MurmurX, a fully local voice-first AI pipeline (WebRTC VAD → Whisper.cpp → Qwen-2.5 3B → TTS) achieving sub-2-second end-to-end latency with zero cloud dependency — demonstrating privacy-first, cost-efficient AI system design. - Built and evaluated ML/time-series models across Ethereum, Bitcoin and Nasdaq market data, combining technical indicators, walk-forward validation and error analysis with applied LLM/RAG project work.	
Software Academy Coding Tutor & Sales Specialist London, UK · Hybrid	Feb 2022 – Jun 2022
- Managed a team of 3 tutors and mentored 50+ students (ages 6–16) in Python and Unreal Engine; consistently among the highest-rated instructors for student satisfaction. - Generated approximately £15,000 in course revenue through consultative client engagement, achieving the highest conversion rate on the team.	

PROJECTS

WorkflowForge – Workflow Intelligence & Automation Engine

2026

Python · Rule-Based NLP · Mermaid Diagrams · CLI · pytest · Ruff · GitHub Pages | github.com/MohamadKanso/WorkflowForge · [Live Demo](#)

- Built an open-source workflow intelligence engine that ingests unstructured operational inputs (meeting notes, SOPs, transcripts, spreadsheets) and outputs structured workflow maps, bottleneck analysis, Mermaid architecture diagrams, and ranked automation opportunities — zero API keys or accounts required.
- Implemented a transparent scoring engine ranking automation opportunities across impact, feasibility, complexity, and operational risk, with a browser-based live demo, Python CLI, pytest test suite, and full contributor documentation.

VoiceSafeKit – Privacy & Safety Layer for Voice Assistants

May 2026

Python · JavaScript · Whisper · MediaRecorder · Transformers.js · pytest · Ruff · GitHub Pages | github.com/MohamadKanso/VoiceSafeKit

- Built an open-source transcript safety system classifying voice assistant inputs as SAFE, REDACT, REVIEW, or BLOCK before they reach an LLM, returning risk scores, plain-English findings, and a sanitised transcript.
- Added browser microphone recording, Whisper-based in-browser transcription, four input modes (record, upload, paste, type), privacy and safety detectors, Python CLI, pytest coverage, a live demo, and an OpenVoiceOS-compatible adapter.

CodeSentry – Local AI Code Review & Auto-Fix Platform

Feb 2026

Python · Ollama · Local LLMs (Llama 3.2, Qwen 2.5) · React · Flask · SQLite · GitHub Actions · pytest | github.com/MohamadKanso/CodeSentry

- Built a fully local AI code-review platform scanning full codebases for bugs and security issues with SQLite result caching for instant re-runs and Markdown report export — zero cloud dependency, full CI/CD integration via GitHub Actions.
- Added Git-aware staged reviews, commit analysis, pre-commit hook installation, patch-based auto-fix, YAML severity configuration, interactive LLM Q&A, and a modular 10-component Python architecture.

Enterprise AI Assistant – MCP/RBAC Tool-Use Prototype

2026

Python · FastAPI · LangChain · MCP-style Tools · RBAC · Keycloak/SSO · PostgreSQL · Redis · React · Pydantic

- Designed and built an end-to-end enterprise-style AI assistant prototype that lets authorised users query structured customer and issue data through approved MCP-style tools rather than unrestricted database access.
- Continuing to evolve the prototype independently toward production readiness, hardening Keycloak/SSO role handling, RBAC checks, PostgreSQL retrieval, Redis trace storage, structured outputs, source/trace UI and human-review controls.

IncidentForge – AI Agent Evaluation & Incident Simulation Toolkit

May 2026

Python · CLI Tooling · Synthetic Data · Logs/Metrics/Traces · SQLite · pytest · Ruff | github.com/MohamadKanso/IncidentForge

- Built an open-source toolkit generating deterministic, production-style incident bundles (alerts, metrics, logs, traces, runbooks, ground truth) across four realistic failure scenarios for benchmarking AI SRE agents.
- Implemented a five-dimension weighted RCA scoring engine (root cause **42%**, evidence **34%**, remediation, service identification, red-herring resistance) with SQLite scoring history, HTML/JSON/Markdown reports, and zero third-party runtime dependencies.

SignalBrief – Autonomous AI Analyst Briefing System

2026

Python · LangGraph · Pydantic · Streamlit · Tavily · Plotly · pytest · Ruff | github.com/MohamadKanso/SignalBrief · [Live Demo](#)

- Developed a **6-agent LangGraph pipeline** (Planner → Search → RAG Reader → Extraction → Risk/Opportunity → Briefing Writer) converting any company or topic into a structured analyst briefing with typed facts, numeric metrics, risk scoring and sentiment analysis.
- Engineered a custom lightweight TF-IDF evidence index (zero vector-DB dependency), Pydantic-typed structured outputs with JSON export, Tavily live search with a deterministic demo fallback so recruiters can run the app with **zero API keys**, per-agent latency tracking, Plotly sentiment gauge, pytest coverage, and Ruff CI.

EDUCATION

City, University of London

Sep 2020 – Aug 2024

MSci Data Science

Upper Second Class Honours (2:1) with First-Class performance in Mathematics and Applied Machine Learning modules.

Key modules: Computer Vision (**91%**), Mathematics for Computing (**88%**), Applied Machine Learning, and Mathematics for Data Science.

Dissertation: Time-Series ML System for Financial Market Data – First-Class mark (**76%**). Built LSTM and regression models on BTC/USD intraday data with walk-forward validation, technical indicator engineering, and rigorous out-of-sample evaluation. github.com/MohamadKanso/Dissertation

ADDITIONAL

Awards: Computer Vision Challenge Winner (2023); Cognizant Vibe Coding Challenge Winner (2025); UKMT Senior Maths Challenge Silver (2020)

Languages: English (Fluent); Arabic (Fluent); French (Conversational)

Interests: Generative AI & LLM systems; AI agents; RAG; workflow automation; open-source engineering; chess (Lichess ~1900 Rapid)