

SHAIBI SHAMSUDEEN

AI SYSTEMS ARCHITECT

Dubai, United Arab Emirates | +971 55 3423631 | shaibis@gmail.com | [Linkedin](#) | [OpenReview](#)

PROFESSIONAL SUMMARY

AI Systems Architect with 10+ years designing intelligent applications, data and machine learning platforms, workflow automation, and governed enterprise systems. Specializes in agentic AI architecture, with deep focus on **loop engineering** (agent control loops, plan-act-observe and tool-use cycles, reflection and evaluator-optimizer loops, state checkpointing, termination criteria, and human-in-the-loop controls) and **prompt engineering** (system-prompt design, context engineering, structured outputs, few-shot and chain-of-thought patterns, prompt versioning, and evaluation). Architects multi-agent orchestration, RAG pipelines, tool-calling interfaces, and LLM evaluation, guardrail, and observability layers that balance latency, token cost, reliability, and auditability. Translates business, compliance, and non-functional requirements into modular, scalable reference architectures from design through deployment. Co-author of peer-reviewed research on LLM safety, mechanistic interpretability, and inference-time refusal control.

CORE EXPERTISE

Loop Engineering: Agentic control loops, ReAct (reason-act-observe) cycles, tool-use loops, reflection and self-correction, evaluator-optimizer loops, retry, fallback, and termination logic, stateful graph orchestration, human-in-the-loop approvals, closed-loop feedback

Prompt Engineering: System and role prompt design, context engineering, few-shot and chain-of-thought prompting, prompt chaining, structured outputs (JSON schema), function and tool calling, prompt templating and versioning, prompt regression testing, prompt-injection defense

AI Systems Architecture: Generative AI, LLM applications, agentic workflows, multi-agent orchestration, RAG, semantic search, vector databases, model routing, caching, token budgeting, LLMops

AI Reliability and Evaluation: LLM evaluation, eval harnesses, backtesting, guardrails, AI safety, refusal behavior, activation steering, observability and tracing, cost, latency, and performance trade-offs

Solution Architecture: System design, target-state architecture, API-first and event-driven design, data architecture, integration architecture, non-functional requirements (scalability, availability, security), requirements engineering

Engineering Stack: Python, TypeScript, SQL, React, Flask, FastAPI, REST APIs, data pipelines, system integration

Data and AI: pandas, NumPy, scikit-learn, TensorFlow, PyTorch, Matplotlib, LangGraph, LangChain, OpenAI, Gemini, Hugging Face, ChromaDB

Platforms and Delivery: PostgreSQL, MySQL, Supabase, AWS, Docker, Git, GitHub, CI/CD, Vercel, Power BI, Tableau, Streamlit

Leadership and Governance: Technical leadership, cross-functional delivery, Agile/Scrum, stakeholder management, GRC and auditability

PROFESSIONAL EXPERIENCE

Head of Engineering Unit | Proceedit Trading | Barcelona, Spain (Remote from Dubai) | Jan 2024 - Present

- Act as lead systems architect for AI-driven trading engines, analytics services, and automation bots, leading a five-person team of three quantitative analysts and two software engineers.
- Own end-to-end solution architecture: translate business requirements into target-state architecture, data models, API contracts, and non-functional requirements (latency, scalability, reliability), and govern delivery through deployment and performance review.
- Architected a modular, layered pipeline architecture on Python, TensorFlow, PostgreSQL, and Flask for market-data ingestion, price prediction, correction, forecasting, and trade analytics, with API-based integration between decoupled services.
- Designed a new combination algorithm within the correction engine, a post-processing layer between stock-price prediction and order and signal calculation; reduced prediction error to below 1%.
- Established architecture standards for component decoupling, model lifecycle management, and closed-loop performance feedback that compares predicted against actual outcomes to drive recalibration.

Data Analyst | Avenir Arabia IT DMCC | Dubai, UAE | May 2023 - Dec 2023

- Translated client process problems into analytics requirements; cleaned and consolidated multi-source data and built Python, Power BI, and Tableau dashboards for decision support.
- Defined KPIs with business teams, quantified automation opportunities, and communicated findings to senior stakeholders.

Head of Product Development and Compliance Services | Tadashie FZCO | Dubai, UAE | Jan 2021 - Apr 2023

- Led functional and product design for GRC applications covering due diligence, conflicts of interest, anti-bribery controls, document management, approvals, and audit evidence.

- Integrated Python dashboards and API-connected workflows, improving process efficiency by 40% and increasing traceability of compliance decisions.

Senior Compliance Advisor and Functional Systems Lead | Petrofac International UAE Ltd. | Sharjah, UAE | Oct 2013 - Apr 2020

- Led functional architecture and product requirements for the Cerico third-party risk assessment platform, defining risk models, due-diligence questionnaires, workflow logic, approval controls, and audit evidence; Dow Jones acquired Cerico in 2018.
- Designed and evolved global compliance-platform workflows for policies, e-learning, gifts and hospitality, conflicts of interest, annual certifications, and risk-based due diligence, including ERP integration.
- Led compliance risk assessments, control reviews, investigations, policy implementation, and stakeholder adoption across multinational operations.

Quality Assurance Coordinator and Technical Assistant | Petrofac International UAE Ltd. | Sharjah, UAE | Jan 2008 - Oct 2013

- Designed an internal audit planning and findings-analysis tool, supported risk-based audit programs, and automated inspection coordination, reducing inspection costs by 30%.
- Managed quality systems, KPI reporting, lessons learned, and engineering and procurement assurance across major FEED and EPC projects.

SELECTED AI AND SYSTEMS PROJECTS

Luminar Multi-Agent AI Deep Researcher | Python, LangGraph, LangChain, ChromaDB, Streamlit | 2025

- Architected a modular multi-agent research system on a LangGraph stateful agent loop that routes work across web, video, academic, news, and vector-retrieval agents; engineered agent prompts and structured outputs, and added source consolidation, contradiction detection, caching, token and cost tracking, and PDF, DOCX, Markdown, and JSON export.
- Improved information-retrieval efficiency by 45% against the project baseline and won Second Prize among 250+ participants in the Outskill AI Engineering Accelerator.

Conversational Real Estate Intelligence Platform | React, TypeScript, Vite, Gemini Live, Google Maps, Supabase | 2025 - 2026

- Built a real-time voice and text experience with a model tool-calling loop, Maps grounding, property discovery, client profiles, persistent favorites, and photorealistic 3D map interaction; deployed for AIN Global Real Estate.
- Extended the platform into a production digital deployment for Lockwood and Carter, a UAE real-estate brokerage, combining modular project content, responsive front-end architecture, AI-assisted customer interaction, analytics, and conversion-focused landing pages.

Platform-Level Child AI Safety and Regulatory Assurance | Product Development (In Development) | 2026 - Present

- Developing the functional architecture for an operator-side age-assurance control and evidence layer that converts verified age signals and policy rules into executed controls, audit evidence, and continuous assurance; in customer-validation stage.

AI-SDA Workbench | Product Development (In Development) | Next.js, TypeScript, Vitest, Playwright | 2026 - Present

- Architecting an open-source, evidence-driven workbench to design, evaluate, and assure AI-enabled applications across a staged lifecycle loop, from concept validation to pre-deployment approval and change management; MVP with Loop 1 live.

RESEARCH AND PUBLICATIONS

Research focus: Interpretable LLM refusal behavior using Llama 3 8B residual-stream activations, category-specific steering vectors, and inference-time control; evaluation of safety and over-refusal behavior across benign, ambiguous, and harmful prompt sets.

- **From Refusal Tokens to Refusal Control: Discovering and Steering Category-Specific Refusal Directions.** Co-author. Accepted, COLM 2026 Actionable Interpretability Workshop; arXiv:2603.13359.
- **What Do Refusal Tokens Learn? Fine-Grained Representations and Evidence for Downstream Steering.** Co-author. Accepted poster, NeurIPS 2025 Mechanistic Interpretability Workshop.

EDUCATION AND PROFESSIONAL DEVELOPMENT

Master's Degree in Data Science | College de Paris | 2024

Master of Science in International Management | University of Liverpool | 2013

Bachelor of Science in Hospitality and Administration | IHMCT, India | 2004

- AI Engineering Accelerator, Outskill, 2025; Second Prize project award.
- Outskill Alumni continuing development, 2025 - Present: 40+ hours of focused labs in OpenClaw, Hermes Agent, Claude Code and Antigravity, Windsurf and Devin, open-source AI, production reliability, AI evaluations, delegation, and problem selection.
- Scrum Fundamentals Certified | IBM Data Analyst Capstone | KPMG Australia Data Analytics Consulting Virtual Internship | CCEP-I | ISO 37001 and ISO 9001 auditor training.

Earlier experience: Damac Properties, Unigreen Shipping and Air Freight, and Convergys India Services, 2004 - 2008.