SHAIBI SHAMSUDEEN AI SYSTEMS PORTFOLIO

AI architecture, engineering leadership, research, and enterprise systems

Dubai, United Arab Emirates | shaibis@gmail.com | LinkedIn | GitHub | OpenReview

POSITIONING

I design AI systems that connect models, data, tools, controls, and user workflows. My work spans multi-agent research, retrieval-augmented generation, conversational and spatial AI, machine-learning pipelines, model evaluation, compliance automation, and enterprise risk systems. I combine hands-on engineering with functional architecture: clarifying the business problem, defining system behavior and controls, selecting components, guiding delivery, and validating outcomes.

ARCHITECTURE CAPABILITIES

Capability	Methods	Evidence
Agentic AI	Multi-agent orchestration, tool calling, routing, state, caching	Luminar; Gemini Live real-estate platform
RAG and knowledge	Embeddings, vector retrieval, source synthesis, contradiction checks	Luminar with ChromaDB and multi-source research
AI reliability	Evaluation sets, backtesting, refusal and over-refusal analysis	Algoverse research; Proceedit trading systems
Enterprise integration	APIs, data pipelines, workflow and ERP integration	Cerico, Petrofac compliance platform, Supabase applications
Governance by design	Risk logic, approvals, audit evidence, policy and control workflows	Cerico; GRC applications at Petrofac and Tadashie

SELECTED OUTCOMES

- Second Prize among 250+ participants for Luminar in the 2025 Outskill AI Engineering Accelerator.
- Reduced price-prediction error from 4.5% to 1.3% through model iteration and structured backtesting.
- Improved compliance workflow efficiency by 40% through automation and integrated GRC design.
- Reduced inspection costs by 30% through centralized coordination and process automation.
- Co-authored LLM-safety research presented at a NeurIPS 2025 workshop and extended into a 2026 arXiv preprint.
- Led functional architecture for Cerico, a third-party risk platform acquired by Dow Jones in 2018.

01 LUMINAR MULTI AGENT AI DEEP RESEARCHER

Role Solution architect and lead builder

Problem Knowledge workers need to investigate complex questions across web, video, academic, and news sources without manually stitching together disconnected results.

Architecture A Python and Streamlit application orchestrated with LangGraph and LangChain. Specialized agents query distinct source types in parallel; a consolidation layer merges findings and detects contradictions. ChromaDB supports vector retrieval, while caching, validation, rate handling, token tracking, and cost estimation support operational use. Reports export to PDF, DOCX, Markdown, and JSON.

Contribution Defined the agent topology, state flow, prompts, service abstractions, retrieval pipeline, result model, export layer, and user experience. Integrated OpenAI and multiple research APIs and organized the codebase into agents, services, workflows, models, prompts, and utilities.

Outcome The project reported a 45% improvement in information-retrieval efficiency against its baseline and won Second Prize among 250+ accelerator participants.

Evidence GitHub

02 LLM SAFETY AND REFUSAL CONTROL RESEARCH

Role AI researcher and co-author

Problem Safety-tuned language models must refuse harmful requests without unnecessarily rejecting benign or ambiguous prompts. Coarse refusal behavior limits model usefulness and makes deployment controls difficult to tune.

Architecture Research used a refusal-token-tuned Llama 3 8B model, residual-stream activation analysis, category-specific steering vectors, lightweight probes, and a whitened low-rank combination for inference-time intervention. Evaluation compared refusal and over-refusal behavior across safety benchmarks and tested transfer to same-architecture variants.

Contribution Contributed to the collaborative research, evaluation, interpretation, and publication process through Algoverse AI Research.

Outcome The initial work was presented as a poster at the NeurIPS 2025 Mechanistic Interpretability Workshop. The extended 2026 paper is available as arXiv:2603.13359 and remains under review.

Evidence arXiv | Project code | Workshop listing

03 CONVERSATIONAL REAL ESTATE INTELLIGENCE

Role AI application architect and developer

Problem Property buyers need natural, location-aware discovery rather than static search forms and disconnected project pages.

Architecture A React and TypeScript single-page application connects Gemini Live for low-latency voice and text interaction, model function calling, Google Maps grounding, Places data, photorealistic 3D maps, Supabase persistence, and state management with Zustand. The design separates the live-session client, tools, shared state, map controller, audio streaming, and UI components.

Contribution Adapted the underlying conversational mapping pattern to real-estate discovery, client profiles, property context, saved favorites, grounded location questions, and responsive 3D map behavior. Built and deployed the application for AIN Global Real Estate.

Outcome Created a production-facing demonstration of multimodal, tool-using AI for brokerage workflows and location-led property exploration.

Evidence GitHub | AIN Global

04 LOCKWOOD AND CARTER DIGITAL PLATFORM

Role Solution designer and full-stack implementation lead

Problem A UAE brokerage needed a premium digital platform that combined property marketing, project content, lead conversion, and an extensible foundation for AI-assisted engagement.

Architecture A component-based React and TypeScript front end built with Vite, reusable content sections, responsive project experiences, and Gemini service integration. Later design work extended the platform with richer interactive presentation and campaign-specific landing pages.

Contribution Translated brand and commercial goals into information architecture, reusable components, project pages, performance-conscious assets, and deployment-ready releases.

Outcome Launched a production website for Lockwood and Carter and established a reusable platform for property launches and ongoing content development.

Evidence GitHub | Live site

05 CERICO THIRD PARTY RISK ASSESSMENT PLATFORM

Role Functional architecture lead

Problem Multinational organizations needed a consistent, auditable way to assess suppliers, clients, and other third parties for bribery, corruption, and broader compliance risk.

Architecture A cloud-based due-diligence and risk-assessment platform combining configurable questionnaires, risk scoring, approval workflows, escalations, evidence capture, registers, and audit trails.

Contribution Led the functional architecture: translated compliance policy into risk models, questionnaire logic, workflow states, approval and escalation rules, control evidence, reporting requirements, and operational use cases. Connected business, compliance, and technology stakeholders during design and implementation.

Outcome The platform supported systematic third-party due diligence and regulatory compliance. Dow Jones acquired Cerico in March 2018, validating the product's commercial and strategic value.

Evidence Dow Jones acquisition

06 ALGORITHMIC TRADING AND ANALYTICS SYSTEMS

Role Head of Engineering Unit

Problem Trading teams needed dependable pipelines and validation methods for predictions, price corrections, forecasting, and decision-support analytics.

Architecture Python, TensorFlow, PostgreSQL, and Flask services process market data, train and serve machine-learning models, and support analytics engines and automation bots. Structured backtesting and defined metrics govern model iteration.

Contribution Lead three quantitative analysts and two software engineers; set architecture direction, convert business objectives into delivery plans, guide implementation, review results, and communicate risk and progress to leadership.

Outcome Reduced reported prediction error from 4.5% to 1.3% while establishing a clearer evaluation and delivery discipline for AI-driven trading products.

07 ALETHEIA AI SAFETY AND ASSURANCE INITIATIVE

Role Independent product researcher and functional systems architect

Problem Youth-facing platforms face increasing obligations to demonstrate that age-related protections are correctly selected, executed, evidenced, and kept effective over time. Existing age-verification providers can supply signals, but operators still need a control and assurance layer around those signals.

Architecture The planned design receives age evidence from established providers and platform-native sources, records provenance and assurance level, evaluates customer-configured rules, triggers the required platform control, captures proof of execution and effectiveness, and monitors freshness, expiry, revocation, conflicting signals, circumvention, and re-check conditions.

Contribution Defined the initial product hypothesis, assurance-state model, control lifecycle, evidence fields, operating boundaries, and staged validation plan. The concept is being developed in a personal capacity with emphasis on youth safety, auditability, privacy, and operator accountability.

Outcome The initiative is in research and validation rather than production. Current work is testing the competitive gap and customer willingness to pay before authorizing a build. The portfolio teaser intentionally presents the system concept without claiming a completed product.

CONTINUING AI ENGINEERING DEVELOPMENT

Completed the Outskill AI Engineering Accelerator in 2025 and remain active in its alumni learning community. Since the accelerator, completed more than 40 hours of focused sessions covering OpenClaw, the Hermes Agent, Claude Code and Antigravity, Windsurf and Devin, open-source AI workflows, moving AI-assisted code into production, AI evaluations and reliability, agent delegation, and problem selection. Ongoing weekly study tracks model releases, agent frameworks, evaluation methods, and production practices.