ATHIRA ANIL KUMAR

AI Engineer | Generative AI | Agentic AI Systems | LLM Security | AI Red Teaming

Sharjah, UAE | +971 5454 9 2345 | ara.ceh24@gmail.com | github.com/ara-5 | linkedin.com/in/athira-a-k | ara-5.github.io/Portfolio | UAE Resident | Available Immediately

PROFESSIONAL SUMMARY

AI Engineer specialising in Generative AI, Agentic AI Systems, RAG, and LLM Security. MSc Cyber Security (Distinction) and IEEE-published researcher. Experienced in building production-ready multi-agent systems, LLM security evaluation frameworks, and scalable automation pipelines using LangGraph, FastAPI, and Docker.

TECHNICAL SKILLS

Generative AI:	RAG, LangGraph, LangChain, Agentic AI, Prompt Engineering, Ollama, FAISS, BM25, Corrective RAG, RAGAs, Cross-Encoder Reranking
AI Security:	Prompt Injection, Jailbreak Testing, System Prompt Extraction, AI Red Teaming, LLM Security Evaluation, Adversarial Prompting, Severity Classification
ML & NLP:	Transformers, DistilBERT, Scikit-Learn, SHAP, LIME, NLP Pipelines, Feature Engineering
RL:	PPO, Q-Learning, Stable Baselines3, Gymnasium, Adaptive Difficulty Adjustment
Backend:	FastAPI, Docker, GitHub Actions, CI/CD, Azure Container Apps, SSE Streaming, REST APIs, n8n
Languages:	Python, JavaScript, C, C++, GDScript

RESEARCH & PUBLICATION

IEEE Conference Paper Published 2025

Detecting Human-Written and AI-Generated Phishing Emails Using DistilBERT and Explainable AI

- Tri-class transformer NLP framework distinguishing legitimate, human-written phishing, and AI-generated phishing emails; SHAP + LIME explainability; approx. 98% accuracy across multiple open-source datasets.

ieeexplore.ieee.org/document/11407086

PROFESSIONAL EXPERIENCE

Gen-AI Security Testing Developer May 2026 – Present

Cosmic Info Solution

- Evaluated AI systems against prompt injection, jailbreak, system prompt extraction, and role confusion attacks; classified vulnerabilities by severity and wrote structured remediation reports.
- Contributed to AI red teaming research and development of the Gen-AI Prompt Injection Tester for assessing chatbot resilience.

Tech: Python, LLM Security, Prompt Engineering, AI Red Teaming, AI Safety Evaluation

AI-Native Developer / Agentic Product Builder May – Jun 2026

- Worked on architecture, agentic AI workflow design, and automation pipelines for an AI-powered video production platform.

Tech: Generative AI, Agentic AI, Workflow Automation, AI APIs, Distributed Systems

ENGINEERING PROJECTS

Enterprise Agentic RAG Platform

github.com/ara-5/Enterprise-Agentic-RAG-Platform | Python, LangGraph, FAISS, BM25, FastAPI, Docker, RAGAs

- Production-grade RAG with LangGraph agentic loop, hybrid BM25+FAISS retrieval, Reciprocal Rank Fusion, 2-stage cross-encoder reranking, and Corrective RAG with Tavily web fallback.
- RAGAs evaluation pipeline (Faithfulness 0.91, Answer Relevancy 0.88) with CI/CD quality gate that fails builds below 0.70 faithfulness.

Mind-MRI Agent Second Brain — Graph-RAG Memory Core

github.com/ara-5/mind-mri-agent-second-brain | Node.js, Graph-RAG, BFS, TF-IDF, REST API, D3.js, JavaScript

- Zero-dependency offline-first Graph-RAG memory system for autonomous AI agents; BFS graph traversal over Obsidian-style wikilinks reduces LLM context input by 95-99% vs. flat-RAG or full codebase injection.
- Implements TF-IDF semantic search, exponential time-decay relevance scoring, token-budget-aware context assembly, and an interactive D3.js force-directed knowledge graph visualiser.

AI Operations Assistant — Multi-Agent BI System

github.com/ara-5/ai-operations-assistant | *Python, LangGraph, Claude Sonnet, FastAPI, Docker, Azure*

- Multi-agent system routing questions through a Manager, Research, Analysis, and Report pipeline with real-time SSE streaming and Azure Container Apps deployment.

Automated Cyber Threat Intelligence Feed

github.com/ara-5/Money-Threat-intel | *Python, Groq/Llama 3, MITRE ATT&CK, GitHub Actions*

- Zero-cost platform aggregating 9 RSS feeds with MITRE ATT&CK tagging, Llama 3 AI briefings, severity scoring, and automated daily distribution to Discord, Slack, and email.

AI-Powered Phishing Detector — IEEE Published

Python, DistilBERT, Transformers, SHAP, LIME

- Tri-class detection (legitimate, human-written, AI-generated) achieving approx. 98% accuracy; dual-threshold strategy reduced false positives. Basis of 2025 IEEE publication.

DungeonRL — Adaptive AI Dungeon Crawler

github.com/ara-5/Dungeon-rl | *Python, PPO, Stable Baselines3, Gymnasium, Pygame*

- PPO RL agent performs real-time Adaptive Difficulty Adjustment by reading gameplay telemetry every 10 seconds and tuning enemy stats to maintain player flow state — 300k-timestep training pipeline.

LifeForge — AI Life Simulation Game

github.com/ara-5/LifeForge | *Godot 4.3, Q-Learning, Gemini Pro, Firebase, AdMob*

- Q-Learning NPCs with real-time Social RL Graph; Gemini Pro LLM narrative generation; dynamic economy via Ornstein-Uhlenbeck and Markov chains; Firebase cloud saves and AdMob monetisation.

AI Content Brief Workflow

github.com/ara-5/AI-Content-Brief-Workflow | *JavaScript, Anthropic API, Prompt Chaining*

- Four-stage chained LLM pipeline (Intent Analysis, Outline, Brief Drafting, Quality Scoring) with XSS-safe output and CSP restricted to api.anthropic.com.

EDUCATION

MSc Cyber Security — Distinction Nov 2025

Modules: Big Data Analytics, Machine Learning for Security, Security Operations & Assurance

B.Tech Electronics & Communication Engineering 2019

LANGUAGES

English (Fluent) | Malayalam (Fluent) | Hindi (Intermediate) | Arabic (Basic)