

Syed Shaan Mahmood

📍 Slough / London, UK | 📞 +44 7466 605683 | 📠 +971 5554 33441

✉️ syedshaanmahmood@gmail.com | 🌐 syedshaan.com | 🌐 syed-shaan-mahmood | 🎮 computator1200

Profile

AWS-certified (×3) final-year Computer Science (AI) student and engineer focused on **AI infrastructure** and the full LLM lifecycle. Currently building AI-assistant infrastructure on placement at BlockAPT, and I run my own GPU/inference home-lab. I communicate clearly, collaborate well across distributed teams, and enjoy owning hard problems end-to-end.

Technical Skills

AI / ML: LLMs & agents (MCP, LiteLLM, Ollama, Hermes, ACP), RAG (Pinecone), PyTorch, TensorFlow/Keras, scikit-learn, OpenAI Whisper, YOLO (CV), prompt engineering, OpenAI & Gemini APIs

Cloud & Infrastructure: AWS (Solutions Architect, AI Practitioner, Cloud Practitioner), GCP, Docker, microservices, Linux, NVIDIA CUDA / GPU acceleration, VPS & self-hosting, Caddy, Nginx, WireGuard, CI/CD

Backend & Data: Python (FastAPI), Node.js (Nest.js), PostgreSQL, Redis, Elasticsearch, Pinecone, REST APIs

Programming & Tools: Python, JavaScript/TypeScript (React, Next.js), Java, C#, C/C++, Rust, Dart, SQL, Bash; Git (GitHub, Bitbucket, Codeberg), Jira, Agile, pytest

Spoken Languages: English (fluent), Japanese (conversational)

Experience

Data Science / AI Engineer — Industrial Placement, BlockAPT

Sep 2025 – Present

Cybersecurity SaaS scale-up — AI Service team

London, UK

- Engineered an **MCP (Model Context Protocol)** server giving the company's AI assistant a comprehensive tool layer over its cybersecurity platform — log ingestion, querying, and automated response actions such as blocking malicious IPs — built as loosely-coupled microservices.
- Re-architected the assistant from a single-vendor integration to a **vendor-agnostic inference layer**, unlocking many model providers (including self-hosted models on local GPU) and removing vendor lock-in.
- Collaborated daily with the distributed **BAP-AI** team across several time zones, adapting how and when I communicated to keep work moving; my pull requests were reviewed and merged into the main codebase by senior engineers.
- Authored and delivered **technical and proposal presentations** to the team and CTO, explaining the motivation, design and phased timelines behind major architectural changes.
- Secured contract with a **major Japanese enterprise client** by stepping in to interpret as PoC in Japanese when the live translation tool failed — helping turn a stalled discussion into a warm and successful business relationship.

IT Department Intern, Masters Group

Sep 2022 – Sep 2023 · Dubai, UAE

- Supported the IT department across hardware, software and networking — troubleshooting issues and delivering practical solutions to business and client needs.

Selected Projects

IRCX — IRCv3 Platform Adapter for Hermes Agent (open source)

Python, asyncio, IRCv3, CI/CD

- Open-source adapter connecting Nous Research's **Hermes LLM agent** to IRC — full IRCv3 stack with SASL (PLAIN / EXTERNAL / SCRAM-256/512), autonomous *observe mode*, and runtime tool-driven channel agency. **91** network-free tests with GitHub Actions CI; live-validated on Libera.Chat and Rizon.

Self-Hosted GPU & Inference Homelab

Linux, NVIDIA CUDA, Docker, WireGuard, Caddy/Nginx

- Run custom and pre-trained models on an RTX 4070 Ti, exposed from a home-GPU node via a **WireGuard** tunnel reverse-proxied (Caddy/Nginx) through a public VPS; built LLM automation pipelines and agentic tooling (OpenClaw, Hermes).

Vindex — Multimedia Indexing & Summarisation Platform

Python, TypeScript, Docker, Elasticsearch

- Semantic search across video/audio via **OpenAI Whisper** transcription and **Gemini 2.5 Flash** summarisation indexed in Elasticsearch; GPU-accelerated, containerised, and architected to scale on cloud GPU. Live at vindex.wf.

AI Music Recommender (Team Lead) — Led a university team building a TensorFlow Recommenders retrieval-and-ranking system; ran weekly delivery and task delegation. github.com/computator1200/Music-recommender

Education

BSc Computer Science with Artificial Intelligence, Queen Mary University of London

2023 – 2027

A-Levels: Mathematics, Physics, Computer Science — **A*AA**

ASCS School, Dubai (2021 – 2022)

Certifications & Awards

- AWS Certified:** AI Practitioner (2025), Solutions Architect – Associate (2025), Cloud Practitioner (2025) · BCG GenAI Job Simulation, Forage (2025)
- 1st place**, Brunel Hack 2.5 blockchain hackathon (2025) — **ZEROx**, a zero-knowledge encrypted password manager on the Internet Computer (also 2nd, Solana/Superteam UK) · **First runner-up**, Curtin University Dubai Hackathon (2023)