

Aryan Golbaghi

SOFTWARE & AI ENGINEER

London, UK | 07496 772002 | aryangolbaghi@gmail.com | [LinkedIn](#) | [GitHub](#)

PROFESSIONAL PROFILE

Software and AI engineer with experience delivering production applications, RAG systems, LLM evaluation pipelines, developer tooling and real-time observability solutions across fintech, developer tools, sports technology and AI startups. Strong in Python, Java, C#, TypeScript/React, AWS, Azure, PyTorch and SQL. First-Class MComp Computer Science (Artificial Intelligence) graduate and first author of an arXiv machine-learning preprint.

TECHNICAL SKILLS

Programming: Python, Java, C#, TypeScript/JavaScript, SQL, KQL

AI / ML: PyTorch, RAG, LLM evaluation and monitoring, model fine-tuning, Langfuse, AutoRAG, DeepEval, OpenAI, n8n

Backend / Data: Spring Boot, SQL Server, MySQL, Neo4j, Qdrant, data pipelines, metadata and data-quality workflows

Cloud / DevOps: AWS, Microsoft Azure, HPC, SLURM, CUDA, DeepSpeed, Amazon Bedrock, Docker, Linux, Mamba, Azure Functions, Blob Storage, GitHub Actions, CircleCI

Engineering: React, ROS, TDD, Agile/Scrum, observability, incident management, real-time systems

Languages: English, Kurdish, Persian

PROFESSIONAL EXPERIENCE

Lendable

Jun 2026 - Present

AI & Infrastructure Analyst | London, UK

- Developed AWS-hosted production applications for macroeconomic analysis, enabling senior management to access decision-support insights.
- Co-designed and developed a retrieval-augmented generation (RAG) system for company-wide knowledge retrieval.
- Built tooling that automated stages of risk analysis, reducing manual work across the workflow.
- Fine-tuned multiple models on company-specific data.

Redgate Software

Jun 2025 - Sep 2025

Software Engineer Intern | Cambridge, UK

- Worked on a command-line utility that generated undo scripts for Flyway migrations, reducing rollback preparation time and improving developer safety.
- Delivered Flyway Desktop features using test-driven development; replaced deprecated third-party UI components and updated associated tests.
- Conducted an AI feasibility study by interviewing all Flyway Desktop tech leads, prioritising use cases, evaluating technical risk and ROI, and presenting recommendations to the team.

SportLight

Oct 2024 - Jun 2026

Part-time Software Engineer | Bicester, UK

- Developed Python features and operational tooling for real-time observability, including metrics, logging, alerting and failure handling, reducing on-site incidents.
- Designed and implemented a real-time failure-management system that cut recovery time by three hours per failure.
- Managed and mentored four monitors supporting live games, guiding effective responses to real-time incidents.
- Collaborated with product and engineering teams to triage issues, ship fixes and write runbooks that prevented repeat faults.

Ockana

Jun 2024 - Sep 2024

AI Software Engineer | London, UK

- Refactored a PDF automation pipeline from a monolithic class into an extensible subclass-based architecture, automating document parsing and processing while reducing manual intervention.
- Built an LLM monitoring and testing framework with Langfuse and AutoRAG; defined evaluation methods to measure output accuracy and improve transparency for stakeholders.
- Developed a synthetic-data generator using DeepEval to evaluate and optimise RAG models.
- Presented LLM solutions and evaluation findings to non-technical stakeholders, translating model behaviour into practical guidance.

RESEARCH & PUBLICATIONS

Enhancing Speech Emotion Recognition via Fine-Tuning Pre-Trained Models and Hyper-Parameter Optimisation

First Author | arXiv, 2025

- Engineered a resource-efficient AutoML/HPO workflow for speech emotion recognition using pre-trained representations, Bayesian optimisation and TPE.
- Achieved 0.97 balanced class accuracy on EmoDB using eight CPU cores; improved cross-lingual zero-shot accuracy by 0.25 on CREMA-D and 0.26 on RAVDESS.

Model Alignment for Reliable LLM Abstention

Reinforcement Learning Lead | Manuscript in preparation

- Leading reinforcement-learning implementation for a multidisciplinary study of LLM uncertainty and refusal mechanisms using GRPO and DPO.
- Fine-tuned pre-trained LLM models on a university HPC cluster using GPU compute, managing batch jobs, resource allocation, training configurations and experiment outputs.
- Preliminary experiments increased recall by 19.2% with GRPO and 83.33% with DPO relative to prior baselines while investigating hallucination reduction and answer coverage.

EDUCATION

University of Sheffield

Sep 2022 - Jun 2026

MComp Computer Science (Artificial Intelligence) - First-Class Honours

Selected results: Computer Vision 89%, Automata 80%, Speech Processing 78%, Robotics 78%, Natural Language Processing 77%.

City and Islington College

Sep 2020 - Aug 2022

A Levels: Mathematics A*, Further Mathematics A, Computer Science B

LEADERSHIP & ADDITIONAL EXPERIENCE

Founding President, Kurdish Society

Sep 2025 - Present

University of Sheffield | Sheffield, UK

- Founded an SU-recognised society, recruited the inaugural committee, grew membership to 30+ in the first month, delivered 5+ events and secured 50 sign-ups at the Activities Fair.

Secretary, Persian Society

Sep 2024 - Jun 2025

University of Sheffield | Sheffield, UK

- Maintained governance records, scheduled committee meetings, produced action-focused minutes and tracked follow-ups to completion.

Kurdish/Persian Interpreter

Mar 2022 - Aug 2024

UK Language Solutions | England

- Interpreted between Kurdish, Persian and English in high-stakes NHS, council, education, immigration and family-court settings, providing accurate communication and cultural mediation.