

Mohammed Mahdi Jahangiri

Forward Deployed Engineer | AI Systems & Integration Engineer

Cardiff/London, UK (Full UK right to work) | mapjiv@live.com | +44 7473 252475

Portfolio | [linkedin.com/in/m-mahdi-jahangiri](https://www.linkedin.com/in/m-mahdi-jahangiri) | github.com/devPirate01 | huggingface.co/devPirate01

AI engineer who **embeds LLMs into existing software** and owns delivery end to end, from ambiguous requirements to live testing with clients and partner teams. MSc in Artificial Intelligence and 1+ year of professional software delivery across **four partner organisations** and, as a solo founder, 5+ real-estate agency clients. Combines hands-on systems integration with on-site client execution to scope ambiguity, ship production middleware, and validate live builds. Fine-tuned a Gemma 4 (2B) that outperformed a Gemma 4 (26B) zero-shot baseline on task routing (**53.4% vs 29.0%**) at sub-second latency.

Technical Skills

Languages: Python, C++, C#, SQL

LLMs & AI: Cloud & local LLM integration, Agent architectures, Structured tool calling, Guardrails, Fine-tuning (QLoRA), Model evaluation, PyTorch, Hugging Face, RAG

Integration & Delivery: FastAPI, REST APIs, Docker, JSON Schema, Telemetry, Git PR, Release management, Debugging

Customer & Delivery: Technical discovery, requirements scoping, stakeholder demos & live test sessions, cross-team coordination, sprint planning (Agile/Scrum), code review, technical documentation

Experience, Projects & Publications

Forward-Deployed R&D Engineer (Contract) — Awkward Studios

Apr 2026 – Jun 2026

Cardiff, UK (Media Cymru)

Forward-deployed R&D and client delivery for CULTVR Lab's 360° dome installation.

- Collaborated across **four partner organisations**, translating evolving stakeholder requirements into delivered software and supporting the project through to a live public QA session
- With no off-the-shelf specification available, scoped the problem under ambiguous requirements and built a custom C# component that corrected 360° dome UI distortion, removing a **key technical blocker**
- Managed contributions through a **GitHub pull-request** workflow and **travelled regularly** to the client venue to test, debug, and validate live builds directly on the physical dome hardware

MSc Dissertation: LLM Agent Middleware & Guardrails

Python, REST APIs, JSON Schema, QLoRA

- Architected **Python agent middleware** connecting an existing C++ application to **cloud-hosted** and local **LLM** endpoints through authenticated REST APIs; implemented structured tool calls for dialogue routing, clue disclosure, and quest-state actions, with **JSON Schema guardrails** blocking invalid actions before execution
- **Fine-tuned** a Gemma 4 (2B) model using **QLoRA** on consumer hardware, increasing task-routing accuracy from **19.0% to 53.4%** and **outperforming a Gemma 4 (26B)** zero-shot baseline (29.0%) at **0.98s latency**
- Built an automated evaluation and telemetry pipeline covering 1,500 trials and **measuring** routing accuracy, output compliance, latency, and token usage; fine-tuning reduced raw clue-policy violations from 25 to 5, while runtime **guardrails** ensured **zero invalid state changes** reached the game
- Open-sourced the full project with **documentation on GitHub** and published model weights on **Hugging Face**

Engineering Team Coordinator — Landell Games

Dec 2024 – Sep 2026

Remote / Sweden

Promoted from Engineering Intern (Dec 2024 – Jan 2025). Technical lead of the engineering team.

- Owned the weekly build and release process, conducted code reviews, and **managed the shared Git repository**
- Led a 3–6 person team across Agile sprints, assigning tasks and raising sprint success from **40% to 90%**
- Profiled and debugged runtime bottlenecks, increasing application frame rate from **10 FPS to 30 FPS**
- Mediated **between engineering and design teams**, resolving **cross-disciplinary** ambiguity into sprint tasks

ChwedlAR: 1st Place (£1,500), DiffTrailblazers — Cardiff University

AR, RAG, Conversational AI

- Served as team lead and lead engineer, coordinating development of an award-winning AR museum prototype
- Built a conversational system for historical characters using **RAG** and character-specific **LLM personas**

Voice and Emotion Integration Prototype

Python, FastAPI, C++, Whisper

- Built a local **FastAPI** service with custom REST endpoints for low-latency bidirectional communication between **Python ML inference** pipelines and an external C++ runtime application
- Integrated Whisper, Wit.ai, and a locally trained Hugging Face classifier to deliver real-time speech transcription, structured intent recognition, and multi-class emotion **classification**
- Streamed structured JSON inference payloads into Unreal Engine 5 behaviour trees for runtime control

Additional Projects & Publication

MetaHomeG: Built a Meta Quest VR app, translating requirements from 5+ real-estate agencies into product features

Meta Quest Store: Shipped a Unity/C# simulation, owning architecture, compliance submission, and store release

ICHORA 2024: Published and presented a **resource-efficient local LLM** dialogue pipeline balancing quality and cost

Education

MSc Artificial Intelligence — Cardiff University

2025–2026

Dissertation: LLM Agent Middleware & Guardrails

- **Computer Vision & NLP:** RF-DETR object localisation; pandas/scikit-learn automated labelling pipelines
- **AI Safety Founding Fellow (Spring 2026):** Completed intensive curriculum covering model evaluation methodologies, catastrophic risk mitigation, and empirical AI governance

BEng Computer Engineering — Azad University, Tehran

2020–2025

- **First-Class Honours;** core modules: Data Structures & Algorithms, C++, SQL & Database Design, Networks