

AHMAD AL MAWLA

AI Engineer | Beirut, Lebanon | 70777592 | ahmadalimawla21@gmail.com
github.com/ahmaddmawla | [Live demo: Ledger AI Agent](#) | [Live demo: Bilingual RAG Assistant](#)

SUMMARY

Computer Science graduate building and shipping production LLM systems. Designed, deployed, and evaluated three AI applications: an autonomous multi-tool agent, a bilingual (Arabic/English) RAG assistant, and a layered LLM evaluation framework. Strong foundation in Python backend engineering and software testing, with hands-on experience in retrieval-augmented generation, agent orchestration, and containerized deployment.

AI ENGINEERING PROJECTS

Ledger — Agentic Financial Research Assistant · [Live Demo](#) · [GitHub](#)

- Built an autonomous AI agent with LangGraph and LangChain that selects and chains five tools — live market data (Finnhub API), RAG retrieval, currency conversion, and sandboxed computation — to answer multi-step financial research questions.
- Implemented **trajectory evaluation** measuring tool-selection accuracy rather than output alone; achieved 5/5 exact-match on a test suite covering single-tool, concept-only, and multi-tool-chaining queries.
- Streamed the agent's full reasoning trace (tool calls, arguments, and results) to the browser over Server-Sent Events for real-time observability.
- Containerized with Docker and deployed to production; secrets injected at runtime, AST-whitelisted expression evaluation to prevent code execution from model-generated input.

Bilingual RAG Customer-Support Assistant · [Live Demo](#) · [GitHub](#)

- Built and deployed a retrieval-augmented (RAG) assistant answering in English, Modern Standard Arabic, and Lebanese dialect, grounded in a knowledge base via ChromaDB semantic retrieval.
- Calibrated a similarity threshold from measured score distributions to separate answerable from off-topic queries, preventing hallucinated prices and policies.
- Engineered multi-turn conversation memory with LLM-based query rewriting so follow-up questions retrieve correct context.
- Tiered models by task (capable model for answers, lightweight model for query rewriting), reducing latency and API cost. Stack: FastAPI, ChromaDB, Gemini embeddings, JavaScript frontend with RTL rendering.

LLM Evaluation Framework · [GitHub](#)

- Built a layered evaluation framework measuring retrieval context recall, answer correctness, groundedness via **LLM-as-judge**, and refusal calibration across a bilingual golden dataset.
- Surfaced and fixed three defects manual testing had missed — a missing knowledge-base entry, a duplicate chunk splitting retrieval, and a grounding gate leak; context recall improved from 83% to 100% and refusal calibration to 100%.
- Produced a unified scorecard writing timestamped JSON baselines, enabling regression comparison across prompt and threshold changes.

TECHNICAL SKILLS

AI / LLM Engineering: Retrieval-Augmented Generation (RAG), LLM agents & tool calling, LangChain, LangGraph, vector databases (ChromaDB), embeddings & semantic search, prompt engineering, LLM evaluation (groundedness, trajectory evals, LLM-as-judge), structured outputs (Pydantic), Google Gemini API, Rasa

Backend & Deployment: Python, FastAPI, REST API design, Docker, Git/GitHub, Render, pytest, SQL / SQLAlchemy, Linux (WSL/Ubuntu)

Testing & Quality: Automated test suites, API testing (Postman), test case design, UAT, bug reporting, Jira

Other Programming: Java, JavaScript, HTML5/CSS, XML/JSON, basic XPath/XSLT

Financial Markets: FX, bonds, forwards, futures, options, interest rate swaps, trade lifecycle, front/middle/back office

PROFESSIONAL EXPERIENCE

AI Engineer Intern — Touch (Lebanon)

Jul 2025 – Aug 2025

- Built a bilingual (Arabic/English) conversational AI assistant using Rasa, integrating REST APIs and external data sources for dynamic responses.
- Tuned the NLU pipeline to improve intent classification and resolve misrouted responses.
- Performed end-to-end testing and debugging of conversation flows to improve response accuracy and reliability.

Web Development Intern — Stories (Why-Stories)

Jan 2026 – Mar 2026

- Conducted UI and functional testing across desktop and mobile, identifying layout and usability issues.
- Ran UAT against Figma functional specifications to validate features before release.

Software Development & QA Intern — Integrated Digital Systems	May 2024 – Jul 2024
- Developed front-end pages with HTML5/CSS and enhanced interactivity using JavaScript, AJAX, and jQuery. - Tested APIs and verified responses during front-end/backend integration. - Managed relational databases with SQL Server, verifying data accuracy and consistency.	
Work-Study Student, Admissions Office — Lebanese American University	Aug 2023 – Jan 2026
Handled confidential student records with accuracy in a client-facing administrative environment.	

EDUCATION

B.S. in Computer Science — Lebanese American University (LAU)	Aug 2022 – May 2026
Member of the Computer Science Club; volunteered in LAU community service initiatives.	

CERTIFICATIONS & LANGUAGES

Certifications: freeCodeCamp — Responsive Web Design; JavaScript Algorithms and Data Structures
Languages: Arabic (Native), English (Fluent)