

Caleb Kan

+44 07828210751 | calebkan1106@gmail.com | linkedin.com/in/caleb-kan | github.com/caleb-kan | calebkan.com

EDUCATION

Imperial College London	London, United Kingdom
Degree: Bachelor of Engineering – BEng, Computing	Sep. 2023 – Jun. 2026
Grade: First-Class Honours (Predicted)	

WORK EXPERIENCE

Software Engineer Intern	Aug. 2024 – Present
Midas Advisory Python	London, United Kingdom
<ul style="list-style-type: none">Architected modular app with open-source LLMs for non-operating expense data processing, integrating web scraping (Playwright, BeautifulSoup4, Selenium) and APIs (FastAPI, Requests) for top 15 US banks; reduced data collection time from hours to minutes and increased accuracy from 85% to 97%.Engineered dual-model system: small LLM for 93% accurate data filtering, large LLM (LLama 3.1) for schema-based reporting; implemented vector database (Milvus) with embeddings, reducing data retrieval time from minutes to seconds.Designed adaptable architecture enabling seamless integration of new models, enhancing company's market intelligence and banking sector analysis capabilities, supporting more informed decision-making across departments.	
Research Intern	Jul. 2023 – Jul. 2023
Department of Computer Science, City University of Hong Kong Python	Kowloon, Hong Kong
<ul style="list-style-type: none">Led development of GDPR compliance checker extension with distinguished professor, utilising BeautifulSoup4, LangChain, and OpenAI's LLM for automated analysis. Achieved 92.5% accuracy in compliance detection across 400+ websites.Designed user-friendly interface displaying website GDPR compliance status, enhancing data privacy awareness. Increased daily active users by 45% and improved user-reported compliance understanding from 3.2 to 4.7 on a 5-point scale.Created one-click feature for automated data erasure requests, streamlining GDPR rights exercise. Increased successful data erasure request rate by 150%, processing over 100 requests within the first month of implementation.	

PROJECTS

ARMv8 Emulator, Assembler, and Visualiser C	May. 2024 – Jun. 2024
<ul style="list-style-type: none">Engineered high-fidelity ARMv8 emulator and two-pass assembler in C, implementing cycle-accurate fetch-decode-execute pipeline, precise memory management, and robust error handling per ARMv8 specifications.Developed multi-threaded SDL2-based GUI for real-time emulation visualisation, featuring drag-and-drop assembly parsing and dynamic rendering of CPU states, registers, and memory maps.	
Chess AI Kotlin	Dec. 2023 – Jan. 2024
<ul style="list-style-type: none">Engineered Chess AI in Kotlin for pawn race game, implementing game logic, ASCII-based terminal GUI, user interaction system with move validation, and minimax algorithm with alpha-beta pruning for efficient AI decision-making.	
AI Research Agent Python, NoSQL	Jul. 2023 – Aug. 2023
<ul style="list-style-type: none">Engineered AI web research agent using Python, LangChain, and OpenAI's LLM; integrated BeautifulSoup4 for scraping and generating cited summaries, X API for automated content posting, and MongoDB Atlas for secure geolocation tracking.Developed data visualisation pipeline for global AI agent usage, leveraging geospatial analytics to map user engagement and derive actionable insights	

TECHNICAL SKILLS

Languages: Native: English · Mandarin · Cantonese
Programming Languages: C · Haskell · HTML/CSS · Java · Kotlin · Python · SQL
Frameworks: Bootstrap · FastAPI · Firebase · Flask · LangChain · MongoDB · MySQL · PostgreSQL · SQLite
Libraries: Matplotlib · Numpy · OpenCV · Pandas · TensorFlow
Developer Tools: Android Studio · Atom · BitBucket · CLion · Git · GitHub · GitLab · GNU Debugger · Google Colab · Hugging Face · IntelliJ · ~~LaTeX~~ · Ollama · PyCharm · Unix · Valgrind · Visual Studio · VS Code · Xcode