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)

EXPERIENCE

Software Engineer Intern — Aug. 2024 – Present
Midas Advisory | Python — London, United Kingdom

- Partnered with Managing Director to automate non-operating expense data processing for top 15 U.S. banks. Developed app integrating 20+ data sources via open-source LLMs, web scraping, and APIs. Implemented asynchronous programming and Ray for parallel processing, reduced data collection time from hours to minutes and increasing accuracy from 85% to 97%.
- Implemented a Milvus vector database with high-dimensional embeddings using the BAAI/bge-m3 model, reducing data retrieval time from minutes to seconds. Built a scalable architecture designed to easily integrate future machine learning models and data sources, enhancing data processing efficiency.
- Developed a hybrid search mechanism employing Milvus's RRFRanker and engineered a dual-model system: a small LLM achieving 93% data filtering accuracy and a large LLM (Hermes-3-8B-GGUF) for detailed, schema-based reporting. Improved company's market intelligence and analytical capabilities to support informed decision-making across departments.

Research Intern — Jul. 2023 – Jul. 2023
Department of Computer Science, City University of Hong Kong | Python — Kowloon, Hong Kong

- Achieved 92.5% accuracy in GDPR compliance detection across 400+ websites by developing a GDPR compliance checker extension with a distinguished professor, utilising BeautifulSoup4, LangChain, and OpenAI's LLM for automated analysis.
- Increased daily active users by 45% and improved user-reported compliance understanding from 3.2 to 4.7 on a 5-point scale by designing a user-friendly interface displaying website GDPR compliance status, enhancing data privacy awareness.
- Boosted successful data erasure request rate by 150%, processing 100+ requests in the first month by creating a one-click feature for automated data erasure requests, streamlining the exercise of GDPR rights.

PROJECTS

ARMv8 Emulator, Assembler, and Visualiser | C — May. 2024 – Jun. 2024

- Engineered a cycle-accurate ARMv8 emulator and two-pass assembler in C, implementing precise memory management and robust error handling per ARMv8 specifications.
- Developed a 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

- Developed a Kotlin AI for a pawn race chess game, implementing complete game logic, move validation, and an ASCII terminal interface. Optimised AI decision-making using a depth-limited minimax algorithm with alpha-beta pruning.

AI Research Agent | Python, NoSQL — Jul. 2023 – Aug. 2023

- Developed a Python AI research agent using LangChain, OpenAI's LLM, Serper API, BeautifulSoup4, and X API for automated search, web scraping, cited summaries, and content posting. Integrated MongoDB Atlas for geolocation tracking.
- Implemented a Streamlit web app enabling anyone to use the AI research agent, integrating a geospatial visualisation pipeline to map worldwide usage using MongoDB Atlas geolocation data. Analysed engagement patterns of 70+ global users.

HONOURS AND AWARDS

1st Place – G-Research London Coding Challenge 2024 — Sep. 2024 – Sep. 2024
Issued by G-Research

- Achieved First Place among 40+ contestants in the G-Research London Coding Challenge, a 3-hour programming contest where participants compete to achieve the highest score on a given problem.

SKILLS

Languages: **Native**: English · Mandarin · Cantonese
Programming Languages: C · Haskell · HTML/CSS · Java · Kotlin · NoSQL · Python · SQL
Frameworks: Bootstrap · Django · FastAPI · Firebase · Flask · LangChain · Milvus · MongoDB · MySQL · PostgreSQL · SQLite · Streamlit
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