

TANG KEN YI

Tel: +65 86618162 | Email: tangkenyi2001@gmail.com | LinkedIn: <https://www.linkedin.com/in/tang-ken-yi-05629421a/> | Personal Portfolio: <https://tangkenyi2001.github.io/>

EDUCATION

NANYANG TECHNOLOGICAL UNIVERSITY Bachelor of Engineering (Computer Engineering)

Aug 2022 – Present
Class of 2026

- Second Class (Upper) Honours/ Honours (Distinction), Current CGPA: 4.47/5.00

UNIVERSITY OF BIRMINGHAM UG Computer Science and Software Engineering

- First Class Honours, CGPA: 3.83/4.00
- AY 23/24 S2 Student Exchange Programme

EXPERIENCE

SAP

Jan 2025 – Present

AI Engineer Intern

- Deployed Large Language Models (LLMs) for production utilizing Docker and vLLM, reducing inference costs and improving response latency through efficient model serving.
- Conducted finetuning of LLMs with LoRA (Low-Rank Adaptation), optimizing performance for specific use cases.
- Collaborated in the development of a full-stack, end-to-end internal ChatGPT platform, focusing on backend services.
- Developed key backend features using Express.js, including file uploads to S3, LLM chat history persistence, and document-based conversational capabilities.
- Enhanced system reliability by implementing circuit breakers, error handling mechanisms, and service resiliency patterns.
- Spearheaded focus group discussions and enablement sessions to understand user pain points and drive product adoption and presented findings to upper management.

Opulent Marketing PTE LTD

May 2022 – July 2022

Web Developer Intern

- Enhanced the company website with JavaScript and CSS, boosting web traffic by 40% within two months through SEO improvements, including optimized meta tags, content updates, and proper image scaling.
- Ensured mobile responsiveness and implemented media queries to improve mobile search rankings.

PROJECTS

ReviVE

- Developed an "X-to-earn" React + Node.js application, rewarding users with tokens for completing upcycling projects.
- Leveraged LLMs to provide customer service and computer vision to automatically verify whether uploaded images reflect completed upcycling.
- Created a backend database and implemented a social-media-like feed for users to share and browse upcycling activities.
- Achieved 3rd Runner-Up at the EasyA x BCG hackathon for innovation in sustainability technology.

Fake News Detection Agent

- Built an intelligent agent with LangGraph, detecting fake news by scraping the web and performing real-time fact-checking.
- Implemented a local Retrieval-Augmented Generation (RAG) pipeline leveraging Ollama and ChromaDB for claim verification.
- Conducted explainable, context-aware analysis by comparing claims against trusted sources stored in a vector database.

SKILLS & INTERESTS

Languages: Proficient in English and Chinese, conversant in German

Technical Skills: Python, C, Java, SQL, React, JavaScript, TypeScript, Docker, Git

Certificates: Generative AI with Large Language Models by DeepLearning.AI