Ray Zhang

LinkedIn: https://www.linkedin.com/in/ray-zhang-8550ba301/
GitHub: https://github.com/zerayo714
Email: mangoray123@gmail.com
Mobile: +886-913969916

EDUCATION

- **National Sun Yat-sen University (NSYSU)**   Kaohsiung, Taiwan
  Master of Information Management (MIS), AI Laboratory   *Sep 2023 - July 2025(Expected)*
  Courses: Generative AI, Data analysis and visualization

- **National Changhua University of Education (NCUE)**   Changhua, Taiwan
  Bachelor of Information Management(MIS)   *Sep 2019 - June 2023*
  Courses: DS and Algorithm, MIS, Programming, Database, Operating System, Network

EXPERIENCE

- **Winbond – Leveraging AI to Improve Manufacturing Operations**   Taichung, Taiwan
  Software Engineer Intern   *July 2024 - August 2024*
  - **AI Log Intelligence Platform**: Designed and implemented an AI-powered manufacturing log intelligence system by ingesting factory error-reporting APIs into Elasticsearch, performing data normalization, chunking, and vectorization for downstream LLM-based root-cause analysis and troubleshooting.
  - **Multimodal AI Assistant**: Built a production-oriented multimodal AI assistant that processes natural-language queries and image inputs via LangChain-based retrieval pipelines to support manufacturing troubleshooting.
  - **System Automation**: Implemented scheduled data ingestion and indexing pipelines to ensure continuously updated operational knowledge for real-time AI-driven decision support.

  Impact: Automated analysis of approximately 50 daily manufacturing error cases, resolving about 50% via AI and reducing manual engineer investigations by around 25 cases per day.

- **AI Next Mine – LLMOps and AI Platform Engineering**   Remote
  AI Engineer Intern   *September 2025 - January 2026*
  - **Engineering Document & CAD Intelligence Platform**: Built a containerized, API-driven AI platform supporting over 30 document formats and native CAD engineering drawings, enabling real-time visualization, semantic search, and LLM-powered Q&A over structured design data.
  - **Multimodal Retrieval**: Designed an LLM-based pipeline integrating speech (via Whisper transcripts), images, and text into a unified vector database for scalable knowledge retrieval.

  Impact: Enabled web-based visualization and AI-driven analysis of CAD engineering drawings, eliminating the need for native CAD software and enabling natural-language queries over structured design data.

PROJECTS

- **RAG Chatbot**:
  - **LLM-powered Retrieval System**: Built a LangChain-based RAG system with document ingestion, embedding, and vector storage, enabling users to upload files and perform context-aware Q&A via a Streamlit interface.

SKILLS

- **Programming**: Python, Java, C++, C#, TypeScript, SQL

- **Infrastructure & DevOps**: Docker, Microsoft Azure, Git

- **Databases**: MySQL, PostgreSQL, Weaviate

- **AI & Workflow**: LangChain, Elysia, Haystack, n8n

- **Interests**: Generative AI, Smart Manufacturing, Engineering, Investing