

Chia-Hsuan Chang, PhD

✉ shane.chang.tw@gmail.com  [LinkedIn](#)  New Haven, CT

Summary Applied AI scientist specializing in LLMs, natural language processing, retrieval systems, and scientific discovery. Delivered models and agentic systems that generate expert-validated therapeutic hypotheses, reduce LLM costs, and improve drug-representation accuracy.

Experience

Yale University, Department of Biomedical Informatics & Data Science New Haven, CT, USA
Yale-Boehringer Ingelheim Biomedical Data Science Fellow *Jul. 2025 – Present*
Postdoctoral Associate *Nov. 2024 – Present*

- Built CoreELM / [ctELM](#), aligning Llama and Gemma with biomedical vector spaces via adapter training and SFT/LoRA. In blinded physician pairwise evaluations, physicians selected ctELM-generated outputs as real clinical trial abstracts in 44% of pairs.
- Developed an LLM-based hypothesis generation system that identified novel CAR-T therapy hypotheses, subsequently validated by cell biologists.
- Built a GraphRAG agent that identified expert-validated mechanism paths linking antihypertensive drugs to Parkinson's disease.
- Adapted Llama 3.1-8B into a drug text encoder through continual pretraining on 7.15M PubMed abstracts and contrastive alignment of 19K DrugBank descriptions, increasing drug discrimination accuracy from 72.2% to 82.6%.
- Created cost-efficient biomedical literature clustering methods: LITA reduced LLM requests by 80%, while TopicForest used one-third of baseline tokens to produce more interpretable topic hierarchies ([Live demo](#)).

Drexel University, College of Computing and Informatics Philadelphia, PA, USA
Postdoctoral Researcher *Feb. 2023 – Oct. 2024*

- Developed LLM systems for automated cancer staging, clinical data extraction, and medical question answering (QA); multi-agent QA improved accuracy by 6% over single-agent baselines.

Education

National Sun Yat-Sen University Kaohsiung, Taiwan
Ph.D. in Information Systems *Sep. 2016 – Jul. 2021*
Honors: Phi Tau Phi Honorary Member, Exceptional Doctoral Dissertation

National Sun Yat-Sen University Kaohsiung, Taiwan
B.B.A. & M.B.A. in Information Systems *Sep. 2011 – Jun. 2016*

Skills

Software/Infrastructure: Python, R, SQL, Bash, Linux, Docker, Slurm, Git, Neo4j, Milvus, Chroma, MongoDB, SQLite, MySQL

Machine Learning/LLMs: PyTorch, Hugging Face Transformers, LangChain, vLLM, SGLang, Ollama, supervised fine-tuning (SFT), low-rank adaptation (LoRA), multi-modal alignment, contrastive learning, retrieval-augmented generation (RAG), GraphRAG, LLM-as-a-judge

Languages: Mandarin (Native), English (Fluent)

Research Output & Open Source Release

23 peer-reviewed publications, including 15 first-authored papers. Complete list in [Google Scholar Profile](#). Selected open-source repositories: [CoreELM](#), [TopicForest](#), and [StructuredPubMed](#) (Dataset).