LUCAS (HUNG CHUN) HSU

Email: lucas.hsuhungchun.apply@gmail.com Homepage: lucas-hc-hsu.github.io

RESEARCH INTEREST

To develop ***Language-Agent Architectures*** that enable ***Multi-Agent Systems (MAS)*** with multimodal world perception capabilities, scalable agent auto-routing, and multi-source context processing. My research focuses on 1) ***Multimodal LLMs***, 2) ***Agent Orchestration***, and 3) ***Personalized Conversational Information Retrieval***. This vision originated from my early Graph Neural Networks research starting in 2021, where I conceptualized transforming node-to-node social networks into LLM-human networks—a trending research topic we now call ***Human-Agent Collaboration Systems***.

EDUCATION

M.S. in Data Science (Thesis). National Taiwan University *Sep 2021 - Aug 2023*

- Joint Master's Program with Academia Sinica, National Research Academy of Taiwan
- Graduated 1% of the class (GPA: 4.0/4.0) with Academic Excellence Scholarship and Phi-Tau-Phi Scholastic Honor Membership

B.S. in Engineering and System Science. National Tsing Hua University *Sep 2017 - Jan 2021*

- Specialized in Applied Mathematics and Advanced Statistics for Microelectronic Engineering

PUBLICATIONS

Test-Time Scaling Strategies for LLM-based Generative Retrieval in Multimodal Conversational Recommendations
Hung-Chun Hsu, Yuan-Ching Kuo, Chao-Han Huck Yang, Szu-Wei Fu, Hanrong Ye, Hongxu Yin, Yu-Chiang Frank Wang, Ming-Feng Tsai, Chuan-Ju Wang *submitting to ACL 2026*

CFDA@TREC iKAT 2025: Enhancing Personalized Conversational Search through Query Reformulation and Rank Fusion
Yu-Cheng Chang, Guan-Wei Yeo, Quah Eugene, Fan-Jie Shih, Yuan-Ching Kuo, Tsung-En Yu, Hung-Chun Hsu, Ming-Feng Tsai, Chuan-Ju Wang *NIST TREC 2025*

FincGAN: A GAN Framework of Imbalanced Node Classification on Heterogeneous Graph Neural Network
Hung-Chun Hsu, Ting-Le Lin*, Bo-Jun Wu, Ming-Yi Hong, Che Lin, and Chih-Yu Wang* *ICASSP 2024*

FlashGAN: Framework of Localized Node Augmentation via Semi-supervised Learning in Heterogeneous Graphs with Generative Adversarial Networks
Hung-Chun Hsu, Bo-Jun Wu, Ming-Yi Hong, Che Lin, and Chih-Yu Wang *arXiv preprint*

RESEARCH EXPERIENCE

Research Center for Information Technology Innovation, Academia Sinica *Apr 2024 - Present*
Senior Research Assistant and Project Lead at the CFDA Lab *Taipei, Taiwan*

- Leading the Information Retrieval (IR) research project in collaboration with NVIDIA Research (Taiwan and US) under the supervision of Prof. Chuan-Ju Wang and Prof. Ming-Feng Tsai. Focused on conversational retrieval and Multimodal-LLMs.
- Participated twice in NIST's Text Retrieval Conference (TREC) tracks starting from 2024: Product Recommendations and Interactive Knowledge Assistance (iKAT). Led and coordinated the competition teams with up to 6 direct reports.

Center of Digital, Data and Technology, Cathay Financial Holdings Co. Ltd. *Feb 2024 - Apr 2024*
Data Scientist Intern - R&D *Taipei, Taiwan*

- Developed an enterprise-level Chinese legal retrieval pipeline leveraging LangChain, increasing recall@50 performance from 23% to 39% compared to LangChain's basic two-staged retrieval settings through a novel legal articles pre-classification approach.

Research Center for Information Technology Innovation, Academia Sinica *Sep 2021 - Feb 2024*
Full-time (2023 - 2024) and Part-time (2021 - 2023) Research Assistant at the SNAC Lab *Taipei, Taiwan*

- Developed novel GAN-based data augmentation frameworks for heterogeneous graphs, addressing critical node class imbalance issues in graph neural networks.
- Achieved 14.4% improvement in F-score and 12.3% in PR-AUC over SOTA baseline GraphSMOTE through innovative handling of structured graph $G=(V,E,X)$. Published research in IEEE ICASSP with ongoing follow-up studies.

Representative Project

Advancing NVIDIA X-VILA via Constrained Decoding - **with NVIDIA Research US** *Sep 2025 - Present*

- Enhance NVIDIA's multimodal LLM X-VILA with multimodal object retrieval capabilities (e.g., video retrieval) to ensure reliable multimodal information acquisition and eliminate multimodal hallucinations.

2025 TREC Interactive Knowledge Assistance Track (iKAT) *Jun 2025 - Sep 2025*

- Led a team of 7 members to participate in the iKAT track, developing pipelines for traditional offline conversational search and exploring real-time conversational retrieval methods requiring low latency.
- Achieved 2nd place in Offline Passage Ranking and 3rd place in Online Interactive Generation among participating teams, outperforming the competition median by 10 pts in nDCG@3 and 11 pts in average human evaluation score respectively.

Multimodal Conversational Product Search with MLLMs - **with NVIDIA Research Taiwan** *Feb 2024 - July 2025*

- Developed product recommendation system supporting any-to-any multimodal data using novel MLLM-based generative retrieval to predict consumer purchases, enhanced by our test-time reranking mechanism with average improvements of 14.5 points MRR and 10.6 points nDCG@1 across 3 benchmarks. Research findings are targeted for submission to **ACL 2026**.

2024 TREC Product Search and Recommendations Track *Jun 2024 - Sep 2024*

- Led a 4-member team in the product search and recommendations track. Developed a multimodal product retrieval system leveraging BLIP, ViLT, and BEiT-3 as dual encoders and cross encoders.
- Implemented Weighted Sum Fusion to combine unimodal and multimodal retrieval results, achieving nDCG@10 of 76% compared to 72% from traditional BM25 and SPLADE retrievers.

Insurance Claim Fraud Prediction - **with Cathay Life Insurance Co., Ltd. & National Taiwan University** *Sep 2021 - Sep 2022*

- Designed and implemented heterogeneous social network graphs for large-scale networks and developed data augmentation framework to enhance graph neural network performance in detecting hard-to-identify minority fraud risk accounts.

Academic Services

Complimentary AI Solution Consulting Services *AI Cooperative, Academia Sinica, Taiwan*

- Provided AI technical consulting services to three domestic academic research teams and industry companies across diverse applications: (1) sea turtle blood leukocyte identification, (2) dolphin spectrogram recognition, (3) brain tumor cortical generation, and (4) stock market index prediction using language models for financial text analysis.

Academic Peer Review

- Served as reviewer for 15 papers submitted to conferences including ACL[25], AAAI[25,26], NeurIPS[24,25], ICLR[24, 25], etc.

Talks

All talks presented at Joint Laboratory Seminar, Research Center of Information Technology Innovation, Academia Sinica, Taiwan

The Future of AI Interaction: True Multimodal Conversations *Sep 2025*

- Delivered comprehensive overview of advanced 2025 multimodal retrieval methodologies, demonstrated novel cross-modality alignment loss functions for autoregressive models and proposed future directions in unified retrieval and generation MLLMs.

Generative Retrieval with FM-index and Ranking Objectives *Jan 2025*

- Presented Generative Document Retrieval approaches with ranking-optimized objectives, analyzing FM-index implementation for semantic ID-based document representation to achieve improved search time complexity over traditional systems like FAISS.

Discrete Semantic Representations for Dense Passage Retrieval *Aug 2024*

- Introduced approaches using learnable Semantic IDs for document retrieval, including DSI, VQ-VAE, RQ-VAE, and LMIndexer.

Honors and Awards

2023 Scholastic Honor Membership *awarded by The Phi-Tau-Phi Scholastic Honor Society of the R.O.C. (Taiwan)*

2023 Graduation Ceremony Representative
awarded by the Data Science Master Program, National Taiwan University and Academia Sinica

2023 Academic Excellence Scholarship *awarded by National Taiwan University and Academia Sinica*

2022 2nd Place in National Data Competition - Carrefour Taiwan E-commerce Next Purchase Prediction
awarded by Institute for Information Industry, Ministry of Digital Affairs, Executive Yuan of the R.O.C. (Taiwan)