GUANCHEN WU

E-mail: guanchen.wu@emory.edu | Webpage: guanchenwu1015.github.io
Address: 400 Dowman Dr, Suite N410, Atlanta, Georgia, 30307, USA

EDUCATION

Emory University **Atlanta, GA, USA**
Ph.D. in Computer Science *Aug. 2025 – Present*

- Ph.D. Advisor: Dr. Carl Yang.
- **Research Interests:** AI4Health, Large Language Models, Multimodality LLMs, Multi-agent, RAG & Visual RAG.

Emory University **Atlanta, GA, USA**
M.Sc. in Computer Science *Aug. 2023 – May 2025*

- Research advised by Dr. Carl Yang and Dr. Liang Zhao.
- Focused on graph data mining, large language models, and AI for health.

University of California, Irvine **Irvine, CA, USA**
B.Sc. in Computer Science *Sep. 2019 – Jun. 2023*

- Dean's Honor List for six quarters.

RESEARCH EXPERIENCE

Visual RAG for Long-Document Multimodal Question Answering **Atlanta, GA, USA**
Ph.D. Research, Advisor: Dr. Jiayuan Ding; Dr. Carl Yang *Feb. 2026 – May. 2026*

- Designed a retriever-agnostic RAG framework for long-document multimodal QA; +25–30 absolute F1 points across 5 retriever paradigms on MMLongBench-Doc and LongDocURL.
- Built annotation-based LLM reranking with VLM-generated visual captions as the dominant signal, enabling reliable page selection over text, tables, figures, and layout-heavy long documents.
- Manuscript in preparation.

AI4Health and Clinical LLM Research **Atlanta, GA, USA**
Research Assistant, Advisor: Dr. Carl Yang *2024 – Present*

- Conducted research on privacy-protected PHI annotation, medical ontology extension, and EHR-based clinical decision-making with large language models.
- Resulted in publications at MedInfo 2025, BIBM 2025, Agents4Science 2025, and KDD 2026.

Graph Data Mining and Explainable AI Research **Atlanta, GA, USA**
Research Assistant, Advisor: Dr. Liang Zhao *2023 – 2025*

- Conducted research on ontology extension and explanation generation for text-attributed graph learning models.
- Resulted in publications at Frontiers in Big Data 2024 and ACL 2025.

PUBLICATIONS

6. Yuzhang Xie, Keqi Han, Yunpeng Xiao, Hejie Cui, **Guanchen Wu**, Ziyang Zhang, Kai Shu, Jiaying Lu, Xiao Hu, and Carl Yang. *EHRBench: An Automated and Reliable EHR-based Benchmark for Clinical Decision Making with LLMs*. SIG on Knowledge Discovery and Data Mining (SIGKDD), 2026.

5. **Guanchen Wu**, Yuzhang Xie, Huanmei Wu, Zhe He, Hui Shao, Xiao Hu, and Carl Yang. *Utilizing Large Language Models for Zero-Shot Medical Ontology Extension from Clinical Notes*. International Conference

on Bioinformatics and Biomedicine (BIBM), 2025.

4. **Guanchen Wu**, Zuhui Chen, Yuzhang Xie, and Carl Yang. *Towards Automatic Evaluation and Selection of PHI De-identification Models via Multi-Agent Collaboration*. Open Conference of AI Agents for Science (Agents4Science), 2025. **Spotlight**.

3. Bo Pan*, Zhen Xiong*, **Guanchen Wu***, Yifei Zhang, Zheng Zhang, Yuntong Hu, and Liang Zhao. *TAG-Explainer: Narrating Graph Explanations for Text-Attributed Graph Learning Models*. The 63rd Annual Meeting of the Association for Computational Linguistics (ACL), 2025.
*Equal contribution.

2. **Guanchen Wu**, Linzhi Zheng, Han Xie, Zhen Xiang, Jiaying Lu, Darren Liu, Delgersuren Bold, Bo Li, Xiao Hu, and Carl Yang. *Large Language Model Empowered Privacy-Protected Framework for PHI Annotation in Clinical Notes*. The 20th World Congress on Medical and Health Informatics (MedInfo), 2025.

1. **Guanchen Wu**, Chen Ling, Ilana Graetz, and Liang Zhao. *Ontology Extension by Online Clustering With Large Language Model Agents*. Frontiers in Big Data, 2024.

MANUSCRIPTS

Under Submission

2. Mingyang Wei, Dehai Min, Zewen Liu, Yuzhang Xie, **Guanchen Wu**, Carl Yang, Max SY Lau, Qi He, Lu Cheng, and Wei Jin. *EpiQAL: Benchmarking Large Language Models in Epidemiological Question Answering for Enhanced Alignment and Reasoning*. Under submission.

1. **Guanchen Wu**, Zhe Huang, Yuzhang Xie, Runze Yan, Akul Chopra, Deqiang Qiu, Xiao Hu, Fei Wang, Carl Yang. *MIRAGE: Knowledge Graph-Guided Cross-Cohort MRI Synthesis for Alzheimer's Disease Prediction*. Under submission.

Under Preparation

1. **Guanchen Wu**, Jiayuan Ding, and Carl Yang. *Multi-Aspect Page Annotation for Long-Document Multimodal Question Answering*. Manuscript in preparation.

TEACHING EXPERIENCE

- **Teaching Assistant:** CS551, System Programming, Emory University, 2025 Fall.
- **Teaching Assistant:** CS524, Theory of Computing, Emory University, 2026 Spring.

ACADEMIC SERVICE

- **Reviewer:** TNNLS, 2025.
- **Reviewer:** SIGIR, 2025.
- **Reviewer:** TNNLS, 2024.

INDUSTRY EXPERIENCE

Volkswagen (Anhui) Automotive Co., Ltd **Anhui, China**
Data Analyst Intern *2023*

- Utilized Python and SQL to manage and analyze the employee database, identifying patterns in hiring and internal promotions to inform HR strategies.
- Tracked monthly resume submissions for campus and off-campus recruitment and analyzed offer issuance rates by comparing submitted resumes with offers made.
- Provided monthly data insights to the HR department to support recruitment strategy adjustment.