

Yining Wu

(+1) 626 734 3702 | yining.wu@utexas.edu

EDUCATION

University of Texas at Austin

Aug 2025 – May 2027

Master of Science in Information Studies

Austin, TX

- **Coursework:** Machine Learning, Deep Learning, Natural Language Processing, and AI in Health; **GPA:** 4.0/4.0

Tongji University

Sept 2020 – Jun 2025

Bachelor of Engineering in Vehicle Engineering

Shanghai, China

RESEARCH INTERESTS

Multimodal biomedical AI, knowledge distillation and efficient adaptation of foundation models, clinical data modeling, robust evaluation of medical AI systems, and machine learning for complex biomedical data.

PUBLICATIONS & MANUSCRIPTS

- **Wu, Y.,** Du, T., Fang, J., Zhang, C., Admane, S., Ding, Y. (2026). EmoPatient: An Emotion-Directed Patient Simulator for Realistic Palliative Care Communication Training. AMIA 2026 Annual Symposium, November 7-11, 2026, Dallas, TX, USA.
- **Wu, Y.,** DiGiacomo, P., Ding, Y., Brode, W. (2026). Grounded in Consensus, In Step With Emerging Science: A Consensus-Anchored Multi-Corpus Clinical Chatbot for Long COVID. DAIH Workshop at COLM 2026, October 9, 2026, San Francisco, CA, USA.

RESEARCH EXPERIENCE

Data Science & AI Research Intern, Mayo Clinic

Jacksonville, FL | May 2026 – Aug 2026

Project: “LLM-Based MACE Prediction After Liver Transplantation”

- Led the development of an LLM-based pipeline to predict 5-year major adverse cardiovascular events from three years of longitudinal pre-transplant EHR data for 4,534 liver transplant recipients.
- Designed an extractor–scorer architecture that transforms noisy multimodal patient timelines into risk-relevant clinical events and fine-tunes Qwen2.5 models with LoRA for locally deployable risk scoring.
- Directed systematic evaluations across raw EHR, extracted-event, and combined inputs, including modality ablations, multi-agent and rolling-memory variants, and comparisons with traditional machine-learning and deep-learning baselines.

Project: “TrialAgent: A Multi-Agent Framework for Automating EHR-Based Clinical Trial Emulation”

- Built a benchmark to evaluate agentic systems’ ability to translate clinical trial protocols into computable EHR study definitions, including eligibility criteria, cohort construction, and endpoint specification.
- Designed the annotation and adjudication workflow and contributed to the end-to-end agent framework through publication-ready figures, technical documentation, and manuscript writing.

Research Assistant, AI Health Lab, School of Information, UT Austin

Austin, TX | Jan 2026 – Present

Project: “EmoPatient: An Emotion-Directed Patient Simulator”

- Constructed a de-identified cohort of 354 patients with advanced metastatic malignancies from MIMIC-IV and MIMIC-IV-Note using diagnosis-code filtering, SQL-based extraction, and clinically grounded persona construction.
- Designed a closed-loop multi-agent architecture that estimates turn-level emotional states and converts them into structured control signals, enabling dynamic and interpretable simulation of patient responses.
- Led controlled experiments across 20 clinical cases, 60 personality-conditioned patient profiles, and four LLM backbones, including repeated evaluation, robustness analysis, and component-level ablation studies.

Graduate Research Assistant, Dell Medical School, UT Austin

Austin, TX | Oct 2025 – Present

Project: “Multi-Corpus Clinical Decision-Support Chatbot for Long COVID”

- Developed a multi-source retrieval and synthesis system integrating expert-curated guidance, PubMed literature, ClinicalTrials.gov records, and living systematic-review evidence through source-specific processing pipelines.
- Implemented query-aware literature retrieval, entity-based clinical-trial ranking, intervention–outcome matching, consensus-anchored generation, and deterministic citation and provenance post-processing.
- Designed and coordinated a clinician-in-the-loop evaluation with domain experts to assess the factual reliability, clinical utility, completeness, and safety of generated responses.

SKILLS

Programming & Data: Python, SQL, PyTorch, Transformers, scikit-learn, pandas, NumPy, Git, Linux

LLM/VLM Development: Fine-tuning, LoRA, knowledge distillation, multimodal learning, multi-agent systems, RAG

Clinical AI & Evaluation: EHR cohort extraction, clinical NLP, ablation and error analysis, clinician-in-the-loop evaluation