

Gongyao Jiang

jiang.gongyao@gmail.com | zzoay.github.io



EDUCATION

The Hong Kong University of Science and Technology (Guangzhou) Data Science and Analytics, PhD Candidate, supervised by Prof. Qiong Luo.	Jan 2024 - Present Guangzhou / Hong Kong
Tianjin University Computer Technology, Master, supervised by Prof. Meishan Zhang.	Sep 2021 - Jan 2024 Tianjin
Guangdong University of Technology Network Engineering, Bachelor	Sep 2017 - Jun 2021 Guangzhou

PUBLICATIONS

Multi-Modal Content Understanding and Generation	
- (CIKM 2025; 1st Author) Chart-CoCa: Self-Improving Chart Understanding of Vision LMs via Code-Driven Synthesis and Candidate-Conditioned Answering. <ul style="list-style-type: none">A self-improving VLM framework that establishes a data synthesizing loop through code and enhances reasoning via a candidate-conditioned inference mechanism (test-time scaling).	
- (NAACL 2025; 1st Author) JRE-L: Journalist, Reader, and Editor LLMs in the Loop for Science Journalism for the General Audience. <ul style="list-style-type: none">A multi-agent collaborative framework that mimics the professional editorial workflow, integrating a simulated "Reader" feedback loop to iteratively align scientific reporting with public accessibility standards.	
Socially Aware AI & Human-AI Alignment	
- (IUI 2026; 1st Author) ESR-Coach: Leveraging Large Language Models for Training People to Provide Emotionally Supportive Responses. <ul style="list-style-type: none">LLMs for training humans in emotional communication through a "simulation-demonstration-feedback" loop to help humans operationalize effective and flexible interpersonal strategies.	
- (ACL 2024; 5-th Author) ToMBench: Benchmarking Theory of Mind in Large Language Models. <ul style="list-style-type: none">A systematic and multi-granular benchmark that evaluate LLMs' social intelligence and cognitive capability via comprehensive ToM tests.	
Structured Prediction & Parsing	
- (CL 2024; 1st-Student Author) LLM-Assisted Data Augmentation for Chinese Dialogue-Level Dependency Parsing. <ul style="list-style-type: none">A hierarchical data augmentation strategy that leverages LLMs to synthesize training samples with varying degrees of structural invariance.	
- (ACL 2023 Findings; 1st Author) A Pilot Study on Dialogue-Level Dependency Parsing for Chinese. <ul style="list-style-type: none">A foundational corpus for dialogue-level dependency and a signal-driven transfer mechanism that maps syntax to discourse relations.	
Social Media Content Categorization	
- (IP&M 2022; 1st Author) Fake news detection via knowledgeable prompt learning. <ul style="list-style-type: none">A knowledge-enhanced prompt learning framework designed to bridge the objective gap between pre-training and fine-tuning, thereby aligning the text classification task with the models' intrinsic capabilities.	
- (ECAI 2020; 2nd Author) Document and Word Representations Generated by GCN and BERT for Short Text Classification. <ul style="list-style-type: none">A hybrid architecture merging GCN-derived topology features with BERT embeddings, augmenting local semantic representations with global topic information to enhance short-text classification.	
LLMs for Scientific Discovery and Reasoning	
- (Under Review, 1st Author) ChemisTRAG: Table-based Retrieval-Augmented Generation for Chemistry Question Answering. <ul style="list-style-type: none">Propose a table-based RAG framework to solve the representation alignment problem in knowledge retrieval and utilization in chemistry QA.	
- (Under Review, 3rd Author) ChemThinkBench: Evaluating Fine-Grained Chemical Thinking Abilities of Large Language Models. <ul style="list-style-type: none">Construct a comprehensive framework to evaluate (Vision) LLMs' capabilities on chemistry reasoning problems.	

RESEARCH & INTERNSHIP EXPERIENCE

HCI Lab at HKUST Conducted research of LLMs for Humans and Science advised by Prof. Xiaojuan Ma; spearheaded the ESR-Coach system (IUI 2026) for interpersonal skill training and contributed to benchmarking Multimodal LLMs on chemistry reasoning tasks (submitted to ACL 2026).	Sep 2025 - Jan 2026
CoAI Lab at Tsinghua University / Lingxin Intelligence Conducted research on Social & Emotional AI advised by Prof. Minlie Huang; co-developed training pipelines for emotion-aware LLMs and played a key role in constructing the ToMBench evaluation framework (ACL 2024).	Jun 2023 - Nov 2023

SKILLS & EXPERTISE

- Domains & Architectures:** NLP, (Multimodal) LLMs, Human-AI Interaction, Agentic Workflows, Retrieval Systems (RAG).
- Techniques:** Instruction Tuning (SFT), Reinforcement Learning, PEFT (RLHF/PPO), Data Synthesis, Prompt Engineering.
- Tools:** PyTorch, DeepSpeed, verl, vLLM, LangChain, Elasticsearch.

HONORS & AWARDS

National Scholarship	2023
Extraordinary Academic Scholarship	2021
Outstanding Graduate	2021

COMPETITIONS

Kaggle: Feedback Prize - Evaluating Student Writing Silver Medal (3%, 43/2058)	Mar 2022 - Apr 2022
Mathematical Contest in Modeling, MCM 2020 Meritorious Winner (8%, 564/7199)	Feb 2020 - Mar 2020