

Jiatao Li

(+86) 13506052779 | leejames@stu.pku.edu.cn Personal Website: https://leejamesss.github.io

EDUCATION

• Peking University

Beijing, China

B.S. in Data Science | GPA Rank: 1st in Department

Sep 2022 - July 2026 (expected)

RESEARCH EXPERIENCE

• Wangxuan Institute of Computer Technology, Peking University

Feb. 2024 – Present

Undergraduate Researcher (Advisor: Prof. Xiaojun Wan)

Beijing, China

- **Retrieval Augmented LM**: Developed a training-free unsupervised context selection method that jointly optimizes relevance, diversity, and conflict resolution; designed a taxonomy to investigate strategies for utilizing self-generated documents, enabling adaptive, task-specific context generation.
- Socially Aware NLP:
 - * LLM Intrinsic Bias Probe: Designed dataset and experimental framework to probe LLMs' socio-cognitive biases and attitude shifts under feedback across social dimensions.
 - * **AI Reviewer Bias Diagnosis**: Introduced reusable perturbation framework and dataset for aspect-level bias detection in review automation, enabling fine-grained analysis.
 - * **Detector Sociolinguistic Bias**: Built dataset to benchmark sociolinguistic biases across groups in AI text detectors; introduced stats method to quantify/isolate attribute biases, controlling confounders.
 - * **Bias Mitigation via Interpretable Detection**: Developed multi-agent zero-shot system grounded in rhetoric/linguistics for transparent, interpretable AI text detection.

SALT Lab, Stanford University

Mar. 2025 - Present

Undergraduate Researcher (Advisor: Prof. Diyi Yang; through Stanford UGVR Program)

Stanford, CA, USA

• **Human-Agent Collaboration**: Built simulated-user and experiment protocols for scalable human-AI collaboration evaluation; developed plug-and-play layer decoupling task execution and user interaction to turn any autonomous agent into collaborative agent; outperforms baselines on CoGym, TauBench, OSWorld, and WebArena; applied to OS environments for interactive slide creation with real users.

PUBLICATIONS

- **Jiatao Li**, Xinyu Hu, Xunjian Yin, Xiaojun Wan. Evaluating Self-Generated Documents for Enhancing Retrieval-Augmented Generation with Large Language Models. *NAACL* 2025 *Findings*.
- **Jiatao Li**, Xiaojun Wan. Who Writes What: Unveiling the Impact of Author Roles on AI-generated Text Detection. *ACL* 2025.
- Jiatao Li, Yanheng Li, Xinyu Hu, Mingqi Gao, Xiaojun Wan. Aspect-Guided Multi-Level Perturbation Analysis of Large Language Models in Automated Peer Review. *CIKM* 2025.
- Jiatao Li, Xinyu Hu, Xiaojun Wan. SMART-RAG: Selection using Determinantal Matrices for Augmented Retrieval. *Under Review*.
- **Jiatao Li**, Yanheng Li, Xiaojun Wan. Analyzing Cognitive Differences Among Large Language Models through the Lens of Social Worldview. *Under Review*.
- **Jiatao Li**, Mao Ye, Cheng Peng, Xunjian Yin, Xiaojun Wan. AGENT-X: Adaptive Guideline-based Expert Network for Threshold-free AI-generated Text Detection. *Under Review*.

HONORS AND AWARDS

2025
2025
2023
2023
2023
2023
2024
2024
2020

SKILLS, INTERESTS & OTHERS

- Programming Python, C/C++, Java, LATEX, SQL, HTML, CSS, JavaScript, TypeScript, PHP
- Frameworks & Tools PyTorch, TensorFlow, Scikit-learn, LangChain, Docker, Git, React, MySQL
- Languages Chinese (Native); English (Proficient: TOEFL 107, 21st Century Cup National Final Second Prize)
- Interests Basketball, Swimming, Table Tennis, Badminton, Running