Zhongsheng Wang

✉ jkwzs0429@gmail.com · 🔗 LinkedIn · ⚙ GitHub · 🌐 jkwzs.cn · 👤 Google Scholar · ⓘ ORCID

Summary/Objective

I am currently a first-year PhD student in Computer Science at the University of Auckland, with a research focus on constructing controllable Multimodal Large Language Models (MLLMs). My current research focuses on Multimodal Retrieval-Augmented Generation (MRAG) and the development of trustworthy Multi-Agent systems, with a commitment to making practical and impactful contributions to the field of AI.

Education

The University of Auckland   Auckland, New Zealand
Doctor of Philosophy in Computer Science (Supervisors: Prof. Jiamou Liu and Dr. Qian Liu)   *2025.04 - now*

The University of Auckland   Auckland, New Zealand
Master of Data Science (Supervisor: Prof. Jiamou Liu)   *2022.07 - 2024.09*

Southwest University   Chongqing, China
Bachelor of Engineering in Computer Science and Technologies (Advisor: Prof. Li Li)   *2019.09 - 2023.06*

Research Interests

- Controllable Multimodal Large Language Models (MLLMs) Development
- Multimodal Retrieval Augmented Generation (MRAG)
- Self-Evolving Agents

Research Experience

Arabic Children's Storybook Automated Generation   **2025.04 - now**

- Fine-tuned Qwen3 for Arabic text translation and generation, while leveraging the Flux framework to maintain consistent character identities across storybook illustrations.
- Integrated Meta Voice speech synthesis to enrich storybooks with three aligned modalities–text, illustrations, and speech, supporting the reading and educational needs of local Arabic-speaking children.

ASR Models Fine-tuning in Industry-level CRM Systems   **2024.05 - 2024.07**

- Proposes an industry paradigm for automatically building high-quality datasets for speech model fine-tuning
- The model fine-tuned by this solution has been verified feasible and deployed.

Working Experience

AI Engineer   **2025.04 - now**
Bedaia.ai, New Zealand

- Design an algorithm to achieve controllable generation of detailed storybook illustrations for long texts.
- Image consistency generation was also one of my doctoral research topics.

Graduate Teaching Assistant   **2023.07 - now**
The University of Auckland, New Zealand

- Assist lecturers with computer science courses, including answering questions, marking assignments, and tests.
- The courses I provide help with include:
  * COMPSCI 225 Discrete Structures in Mathematics and Computer Science (2025 S2)
  * SOFTENG 282 Software Engineering Theory (2025 S1)
  * COMPSCI 120 Mathematics for Computer Science (2025 S1)
  * COMPSCI 761 Advanced Topics in Artificial Intelligence (2024 S2)
  * COMPSCI 367 Artificial Intelligence (2023 S2, 2025 S2)

AI Engineer   **2024.05 - 2024.09**
Atom Intelligence, Remote

- Develop speech recognition model fine-tuning solution for the retail industry customer management system.
- Propose a solution for the automated construction of fine-tuning datasets for the retail industry.

- Explore the industrial applications of LLM in the retail industry and provide business solutions.

Data/AI Scientist (Summer Intern) 2023.11 - 2024.02

HouGarden Co, Ltd., New Zealand

- Design the Issue Management System product, which is used to build a ChatBot Question-Answering dataset for the real estate industry.
- Complete the fine-tuning of the large model for the English automated translation of the company's official website and deploy it in practice.

Academic Services

Conference Reviewer

The 40th Annual AAAI Conference on Artificial Intelligence (AAAI'26)

The International Conference on Neural Information Processing (ICONIP'[24-25])

The International Joint Conference on Neural Networks (IJCNN'[24-25])

CONFERENCE ON LANGUAGE MODELING (COLM'25)

2025 ACM Multimedia (MM'25)

ARR 2025 May, July, October

Reasoning and Planning for Large Language Models@ICLR'25 (Workshop)

Journal Reviewer

IEEE Transactions on Fuzzy Systems Proof

Talks

Creative Intelligence: Applications of Large Language Models in Data Generation and Reasoning, University of Electronic Science and Technology of China, Chengdu, China, December 2024

Publications

Conference & Journal Papers

* means equal contribution

Lin Z, **Wang Z**, Zhang X, Liu Q, and Liu J. Narratology Meets Text-to-Image: A Survey of Consistency in AI Generated Storybook Illustrations. Accepted by Artificial Intelligence Review

Li M*, **Wang Z***, Li H, and Liu J. R-Debater: Retrieval-Augmented Debate Generation through Argumentative Memory. Accepted by The 25th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2026

Deng S, **Wang Z**, Mao R, Ciprian D G, and Liu J. CLER: Improving Multimodal Financial Reasoning by Cross-MLLM Error Reflection. Accepted by The 40th AAAI Conference on Artificial Intelligence, AAAI 2025

Li X, Ni L, Wang X, Tang X, Li R, Liu J, and **Wang Z**. LLM-based Business Process Models Generation from Textual Descriptions. Accepted by International Joint Conference on Natural Language Processing & Asia-Pacific Chapter of the Association for Computational Linguistics 2025, IJCNLP-AACL 2025

Wang Z, Lin M, Lin Z, Liu Q, and Liu J. CharCom: Composable Identity Control for Multi-Character Story Illustration. Accepted by ACM Multimedia Asia 2025, MMAsia 2025

Wang Z, Wang S, Wang J, Liang Y, Zhang Y, and Liu J. Weak Supervision Techniques towards Enhanced ASR Models in Industry-level CRM Systems. Accepted by The International Conference on Neural Information Processing, ICONIP 2024

Wang Z, Liu J, Bao Q, Rong H, Liu J, and Zhang J. ChatLogic: Integrating Logic Programming with Large Language Models for Multi-Step Reasoning. Accepted by The International Joint Conference on Neural Networks, IJCNN 2024

Qi Q, Ni L, **Wang Z**, Zhang L, Liu J, and Witbrock M. Epic-Level Text Generation with LLM Through Auto-Prompted Reinforcement Learning. Accepted by The International Joint Conference on Neural Networks, IJCNN 2024

Preprint & Workshop Papers

Xiao X, Shen S, Bao Q, Rong H, Liu K, **Wang Z**, and Liu J. CoRA: Optimizing Low-Rank Adaptation with Common Subspace of Large Language Models. Arxiv pre-print