KANGDA WEI

(919) 904-5280 ◇ kangda@tamu.edu ◇ hansonwei666@gmail.com
L.F. Peterson Building, Room 330, 435 Nagle St, College Station, TX 77843
https://weikangda.github.io/kangda.github.io

RESEARCH INTEREST

Natural Language Processing, Large Language Model, LLM Reasoning, LLM Exploratory Thinking, Post-Training, Multi-Modal, Multi-Agent Systems

EDUCATION

Texas A&M University — College Station, TX
Doctor of Philosophy in Computer Science. Advisor: Ruihong Huang. — Aug 2023 - Dec 2027 (Expected)

University of North Carolina at Chapel-Hill — Chapel-Hill, NC
Master of Science in Computer Science. Advisor: Shashank Srivastava. — Aug 2022 - May 2023

University of North Carolina at Chapel-Hill — Chapel-Hill, NC
B.S. in Computer Science, B.S. in Statistics and Operational Research, GPA: 3.812/4.0 — Aug 2019 - May 2022

INDUSTRY EXPERIENCE

Applied Scientist Intern
Amazon AWS — Santa Clara, CA
Incoming Applied Scientist Intern — May 2026 - Aug 2026

Research Scientist Intern
Bosch Center for Artificial Intelligence (BCAI) — Sunnyvale, CA
Multi-Agent Video Understanding, Mentor: Zhengyu Zhou — May 2024 - Aug 2024

RESEARCH EXPERIENCE

Accelerating Mathematical Reasoning Model Training via Diversity-Aware Reward Reweighting

- Developed MMR-GRPO, a training-efficient approach that integrates Maximal Marginal Relevance into Group Relative Policy Optimization to accelerate convergence by prioritizing semantically diverse completions and down-weighting redundant solutions.
- Designed a parameter-free adaptive mechanism that automatically balances diversity and reward quality without hyperparameter tuning, enabling robust performance across different training configurations.
- Achieved 47.9% reduction in training steps and 70.2% reduction in wall-clock training time while maintaining comparable peak performance across three model sizes (1.5B, 7B, 8B), three RL methods (GRPO, DR-GRPO, DAPO), and five mathematical reasoning benchmarks (AIME 2024, AMC 2023, MATH-500, Minerva Math, OlympiadBench).
- Demonstrated minimal computational overhead (1-5% per-step) through vectorized GPU operations for embedding extraction and similarity computation, making the approach practical for resource-constrained settings.
- *Status:* Paper submitted to ACL 2026; code and models to be released upon acceptance.

Mitigating Bias and Improving Reasoning in LLMs via Exploratory Thinking and RL

- Proposed a novel data generation framework that encourages exploratory thinking in LLMs by prompting them to reason over structurally identical, morally ambiguous story pairs with male and female protagonists.
- Designed a guided judgement reconciliation process to uncover and mitigate biased moral reasoning, producing balanced gender-neutral story-judgement pairs.
- Performed SFT and DPO on LLMs, achieving substantial gender bias reduction on WinoBias and GenMO benchmarks without sacrificing performance on MMLU and TruthfulQA.
- *Ongoing:* Extending framework to other forms of social bias (e.g., race, social status)

Multi-Agent Systems with LLMs/VLMs for Nuclear Safety Anlysis and Video Understanding

- Designed and implemented multi-agent systems that leverage LLMs and VLMs for two domains:
  - *Nuclear Safety:* Developed a BFS-based multi-agent framework for automatic nuclear fault tree generation in collaboration with Idaho National Laboratory (INL). Built a large-scale nuclear safety dataset via OCR and expert validation; benchmarked multiple LLMs/VLMs for fault analysis tasks.
  - *Video Understanding:* Proposed an end-to-end multi-agent multi-modal framework for advanced indexing of lecture/presentation-style videos. Combined VLM-based shot detection with multi-agent collaboration to capture rich cross-modal information for scalable downstream use. Conducted systematic evaluations, including intrinsic human evaluation (Amazon Mechanical Turk) and extrinsic QA benchmarks, on both public and enterprise datasets.

Event Understanding and Coreference in Long Documents

- Introduced *LegalCore*, the first dataset in the legal domain annotated with comprehensive event and event coreference information.
- Benchmarked mainstream LLMs on discourse-level event relation extraction and event coreference tasks, showing that both remain challenging for LLMs.
- Provided strong supervised baselines by fine-tuning smaller language models and conducted extensive empirical analysis, highlighting weaknesses of GPT models in hallucination control, transitivity, and long-range dependency capture.

PUBLICATIONS

- **Kangda Wei**, Ruihong Huang. MMR-GRPO: Accelerating GRPO-Style Training through Diversity-Aware Reward Reweighting. *In submission*
- Cheng Zhang, Rajasekhar Kakarla, **Kangda Wei**, Ruihong Huang. ENG-DRB: PDTB-style Discourse Relation Bank on Engineering Tutorial Video Scripts. *Findings of The Asian Federation of Natural Language Processing and The Association for Computational Linguistics 2025*
- **Kangda Wei**, Hasnat Md Abdullah, Ruihong Huang. Mitigating Gender Bias via Fostering Exploratory Thinking in LLMs. *Findings of the Empirical Methods in Natural Language Processing 2025.*
- Abhilekh Borah, Hasnat Md Abdullah, **Kangda Wei**, Ruihong Huang. CliME: Evaluating Multimodal Climate Discourse on Social Media and the Climate Alignment Quotient (CAQ). *NLP4PosImpact Workshop, Association for Computational Linguistics 2025. Best Paper Award.*
- **Kangda Wei**, Xi Shi, Jonathan Tong, Sai Ramana Reddy, Anandhavelu Natarajan, Rajiv Jain, Aparna Garimella, Ruihong Huang. LegalCore: A Dataset for Event Coreference Resolution in Legal Documents. *Findings of Association for Computational Linguistics 2025.*
- **Kangda Wei**, Zhengyu Zhou, Bingqing Wang, Jun Araki, Lukas Lange, Ruihong Huang, Zhe Feng. Multi-Agent Video Understanding for Advanced Indexing of Lecture Videos. *U.S Patent.*
- Hasnat Md Abdullah, Tian Liu, **Kangda Wei**, Shu Kong, Ruihong Huang. UALBench: The First Comprephensive Unusual Activity Localization Benchmark. *Winter Conference on Applications of Computer Vision (WACV) 2025.*
- **Kangda Wei**, Aayush Guatam, Ruihong Huang. Are LLMs Good Annotators for Discourse-level Event Relation Extraction? *Findings of the Empirical Methods in Natural Language Processing 2024.*
- Jiangshu Du, Yibo Wang, Wenting Zhao, ... **Kangda Wei**, et al. LLMs Assist NLP Researchers: Critique Paper (Meta-)Reviewing. *Proceedings of the Empirical Methods in Natural Language Processing 2024.*
- **Kangda Wei**, Dawn Lawrie, Benjamin Van Durme, Yunmo Chen, Orion Weller. When Do Decompositions Help for Machine Reading? *Proceedings of the Empirical Methods in Natural Language Processing 2023*
- **Kangda Wei**, Sayan Ghosh, Rakesh Menon, and Shashank Srivastava. Leveraging Multiple Teachers for Test-Time Adaptation of Language-Guided Classifiers. *Findings of the Empirical Methods in Natural Language Processing 2023*
- **Kangda Wei**, Sayan Ghosh, and Shashank Srivastava. 2022. Compositional Generalization for Kinship Prediction through Data Augmentation. *Proceedings of the 4th Workshop of Narrative Understanding (WNU2022)*, pages 13–19, Seattle, United States. Association for Computational Linguistics.
- *Songhe Wang, ***Kangda Wei**, Lei Lin, Weizi Li. "Spatial-temporal Analysis of COVID-19's Impact on Human Mobility: the Case of the United States," in the *20th and 21st Joint COTA International Conference of Transportation Professionals.* *Co-author: equal contribution

PROFESSIONAL SERVICES

- Reviewer for ACL Rolling Review (ARR) since February 2024 (covering ACL/NAACL/EMNLP/EACL/AACL).
- Reviewer for IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2025.
- Reviewer for The Computer Journey.

TEACHING SERVICES

Graduate Teaching Assistant:

- COMP211 System Fundamentals, UNC-Chapel Hill, Department of Computer Science, Spring 2023
- COMP431 Internet Services and Protocols, UNC-Chapel Hill, Department of Computer Science, Fall 2022

AWARDS

- Texas A&M University CSE Department Travel Grant, 2023, 2024

SKILLS

- **Programming Languages / Frameworks:** Python, Java, JavaScript, C, HTML, CSS, MATLAB, R, PyTorch, JAX, Hugging Face Transformers, Accelerate, Linux, Parallel Model Training

- **Software & Tools:** PyCharm, VS Code, Spyder, Jupyter Notebook, Weights & Biases,, Docker, Slurm, Git, Tableau, LaTeX, Anaconda, RStudio
- **LLM Training & Post-training:** SFT, RLHF, DPO, GRPO, Reward Modeling, Prompt Engineering, Synthetic Data Generation
- **Cloud & Infrastructure:** AWS, Microsoft Azure, Amazon Mechanical Turk
- **Language:** Chinese (native), English (proficient)