Weixiang Yan

+1 (603) 501 9963 | Homepage | Github | Google Scholar | yanweixiang.ywx@gmail.com

Education

University of California, Santa Barbara (UCSB)
Sept. 2023 – Dec. 2024

*Master of Science in **Computer Science***
California, USA

- **Research Interests:** Code Intelligence, Natural Language Processing, Software Automation

Beijing University of Posts and Telecommunications (BUPT)
Sept. 2018 – June 2022

*Bachelor of Science in **Engineering** and Bachelor of Science in **Management***
Beijing, China

- **Awards & Honors:** National Scholarship(2020 & 2021)(Top 1%), WanDatong Scholarship(Top 5%), Merit Student Award, Outstanding Graduate Award, Outstanding Bachelor's Paper

Work Experience

Amazon (Generative Foundational AI)
Dec. 2024 – Now

Applied Scientist II
California, USA

- Focused on hallucination mitigation, RLVR-based reasoning training, and model evaluation. Led end-to-end governance of hallucination issues for the team, covering post-training data cleaning, evaluation framework development, RLVR-based training, and retrieval quality control, reducing the internal model's hallucination rate from 63.19% to 17.82% (about a 72% relative reduction). Additionally, designed and built an agent interaction environment for RL training and am currently developing an Agentic Evaluation system.

Publications & Patents

- **Weixiang Yan**, Haitian Liu, Tengxiao Wu, Qian Chen, Wen Wang, ..., Li Zhu. "ClinicalLab: Aligning Agents for Multi-Departmental Clinical Diagnostics in the Real World" **(NeurIPS 2025)**

- Zhen Zhang, Xuehai He, **Weixiang Yan**, Ao Shen, ..., Xin Eric Wang. "Soft Thinking: Unlocking the Reasoning Potential of LLMs in Continuous Concept Space" **(NeurIPS 2025)**

- Ilgee Hong, Changlong Yu, Liang Qiu, **Weixiang Yan**, ..., Chao Zhang, Tuo Zhao. "Think-RM: Enabling Long-Horizon Reasoning in Generative Reward Models" **(NeurIPS 2025)**

- Chenghanyu Zhang, Zekun Li, ..., **Weixiang Yan**, Qianyu Zhuang. "SpineBench: Benchmarking Multimodal LLMs for Spinal Pathology Analysis" **(ACM MM 2025)**

- Yuchen Tian*, **Weixiang Yan***, Qian Yang, ..., Lei Ma, Dawn Song. "CodeHalu: Investigating Code Hallucinations in LLMs via Execution-based Verification" **(AAAI 2025)**

- Yuwei Zhao, Ziyang Luo, Yuchen Tian, Hongzhan Lin, **Weixiang Yan**, Jing Ma. "CodeJudge-Eval: Can Large Language Models be Good Judges in Code Understanding?" **(COLING 2025)**

- Qian Yang, **Weixiang Yan**, Aishwarya Agrawal. "Decompose and Compare Consistency: Measuring VLMs' Answer Reliability via Task-Decomposition Consistency Comparison" **(EMNLP 2024)**

- **Weixiang Yan**, Haitian Liu, Yunkun Wang, Yunzhe Li, ..., Hari Sundaram. "CodeScope: An Execution-based Multilingual Multitask Multidimensional Benchmark for Evaluating LLMs on Code Understanding and Generation" **(ACL 2024)**

- **Weixiang Yan**, Yuchen Tian, Yunzhe Li, Qian Chen, Wen Wang. "CodeTransOcean: A Comprehensive Multilingual Dataset for Code Translation" **(EMNLP 2023)**

- Yunzhe Li, Qian Chen, **Weixiang Yan**, ..., Hari Sundaram. "Advancing Precise Outline-Conditioned Text Generation with Task Duality and Explicit Outline Control" **(EACL 2024)**

- **Weixiang Yan**, Qian Chen, Wen Wang, Qianlin Zhang. "DialogueAttention: A training method and device for generating intelligent summarization of multi-person and long conversational text" **(Chinese patent CN 116186244 A)**

- **Weixiang Yan**, Yuanchun Li. "WhyGen: Explaining ML-powered Code Generation by Referring to Training Examples" **(ICSE 2022 Demo Track)**
- Shaokun Zhang, Yuanchun Li, **Weixiang Yan**, Yao Guo, Xiangqun Chen. "Dependency-aware Form Understanding" **(ISSRE 2021)**
- Qian Yang, **Weixiang Yan**, Aishwarya Agrawal. "Investigating Reliable Question Decomposition for Vision-Language Tasks" **(CVPR 2024 Workshop)**
- Qian Yang, **Weixiang Yan**, Aishwarya Agrawal. "Enhancing Multi-Agent Multi-Modal Collaboration with Fine-Grained Reward Modeling" **(NeurIPS 2024 Workshop)**
- Zheye Deng, **Weixiang Yan**, Changlong Yu, Jiashu Wang. "AlphaQuanter: An End-to-End Tool-Orchestrated Agentic Reinforcement Learning Framework for Stock Trading" **(Under review)**

Internship Experience

Alibaba Group (Damo Academy) *July 2022 – Jan. 2023*
Research Intern, advised by Dr. Wen Wang, Qian Chen

- Contributed to several published projects on code intelligence and text generation, including **CodeScope** (ACL), a multilingual executable benchmark for code evaluation; **CodeTransOcean** (EMNLP), a large-scale dataset for multilingual code translation; and a **summarization-enhanced generation** method (EACL). Hold a patent for **DialogueAttention**, improving summarization in multi-person and long-term conversations.

Microsoft Research (MSRA) *June 2021 – Oct. 2021*
Research Intern, advised by Dr. Yuanchun Li

- Designed **WhyGen**, a tool that explains generated code by referring to training samples during code generation and notifies the user of possible recitations and imitations; the paper is published in ICSE.
- Awarded the Honor of Microsoft Research Asia Outstanding Intern "Star of Tomorrow".(Top 10%)

Peking University *Nov. 2020 – May 2021*
Research Assistant, advised by Dr. Yao Guo

- Proposed the **DependEX** method to accurately understand dependencies within forms; the paper is published in ISSRE.

Professional Services

Area Chair: ACL, ARR, AAAI
Conference Reviewer: NeurIPS, ICLR, ACL, EMNLP, NAACL, AAAI, ARR, EACL, COLM
Open Source Services: BigCode