Guangya (Wayne) Wan

Ph.D. Candidate in Data Science, University of Virginia
wxr9et@virginia.edu — (857) 207-5945
gwan.netlify.app — Google Scholar — GitHub — LinkedIn

Research Interests

Large language model reasoning, calibration, and post-training; LLM-based agentic systems and multi-agent coordination; Agentic reinforcement learning; Trustworthy and efficient AI inference.

Education

University of Virginia Charlottesville, VA
Ph.D. in Data Science — Advisor: Prof. Sheng Li Aug 2023 – May 2027 (expected)
Focus: LLM Reasoning, LLM Agents, Trustworthy AI

Harvard University Cambridge, MA
M.S. in Biostatistics (Data Science Concentration) — GPA: 3.89/4.0 Aug 2021 – May 2023

University of Illinois at Urbana-Champaign Urbana, IL
B.S. in Statistics, *Summa Cum Laude* — GPA: 3.94/4.0 Sep 2017 – Dec 2020
Minor: Computer Science & Mathematics

Research & Industry Experience

Google Research – Student Researcher Mountain View, CA — May 2025 – Present
- Developing methods for DeepResearch Agent in Google Cloud Vertex AI Agents Team.
- Conducting research on long-horizon agent reasoning with evolving contextual understanding.

University of Virginia – Graduate Research Assistant Charlottesville, VA — Aug 2023 – Present
- Designed Bayesian optimal stopping frameworks for cost-efficient LLM inference.
- Designed goal-oriented multi-agent framework for psychiatric diagnosis through proactive reasoning and structured knowledge-guided memory.
- Investigated fairness in LLM reasoning across sociolinguistic varieties (**COLM 2025** Workshop).
- Developed novel sampling methods for efficient LLM reasoning (**NAACL, ACL** Findings 2025).
- Published survey connecting causal discovery and large language models (**IJCAI 2025)**.

Chewy – Data Scientist Intern Boston, MA — Jun 2022 – Aug 2022
- Improved demand forecasting accuracy by 20% using time-series clustering (tslearn/DTW) and LightGBM-KNN ensemble models on AWS SageMaker.

Boston Children's Hospital (CHIP) – ML Research Intern Boston, MA — Jan 2022 – Apr 2023
- Assessed fairness of clinical ML models for ED admissions using demographic parity and equalized odds.
- Developed Python package for fairness-aware machine learning on electronic health records.

NCSA (Student Pushing Innovation) – ML Research Intern Champaign, IL — Jun 2019 – Dec 2020
- Built scalable ML pipelines on HPC infrastructure for epidemiological outbreak prediction using 10GB+ datasets with Dockerized workflows and CUDA acceleration.

Technical Skills

LLMs & Deep Learning: PyTorch, Trl (Accelerate), LangChain/Graph, Verl, vLLM, DeepSpeed, ADK
Big Data & Cloud: AWS, Google Cloud, Snowflake, PySpark, SQL, Pandas, Seaborn
Development Tools: Docker, Kubernetes, Claude Code, Slurm, Git, Unix, VSCode, Wandb

Selected Publications & Preprints

1. **G. Wan**, M. Ling, X. Ren, R. Han, S. Li, Z. Zhang.
   "COMPASS: Enhancing Agent Long-Horizon Reasoning with Evolving Context."
   Under review, 2025.

2. **G. Wan**, Z.S. Xu, S. Zorc, M. Baucells, M. Hu, H. Wang, S. Li.
   "BEACON: Bayesian Optimal Stopping for Efficient LLM Sampling."
   Under review, 2025.

3. **G. Wan**, Y. Wu, J. Chen, S. Li.
   "Reasoning-Aware Self-Consistency: Leveraging Reasoning Paths for Efficient LLM Sampling."
   NAACL 2025. arXiv:2408.17017

4. **G. Wan**, Y. Wu, H. Wang, S. Zhao, J. Chen, S. Li.
   "Derailer-Rerailer: Adaptive Verification for Efficient and Reliable Language Model Reasoning."
   ACL Findings 2025. arXiv:2408.13940

5. **G. Wan**, Y. Lu, Y. Wu, M. Hu, S. Li.
   "Bridging Causal Discovery and Large Language Models: A Survey."
   IJCAI 2025. arXiv:2402.11068

6. Y. Wu*, **G. Wan***, J. Li, S. Zhao, L. Ma, T. Ye, I. Pop, Y. Zhang, J. Chen.
   "ProAI: Proactive Multi-Agent Conversational AI for Psychiatric Diagnosis."
   Under review, 2025.

7. R. Zhou*, **G. Wan***, S. Gabriel, S. Li, A. Gates, M. Sap, T. Hartvigsen.
   "Disparities in LLM Reasoning Accuracy and Explanations: A Case Study on African American English."
   COLM 2025 MELT Workshop.

8. X. Li, M. Gao, Y. Hao, T. Li, **G. Wan**, Z. Wang, Y. Wang, X. Chen.
   "MedGUIDE: Benchmarking Clinical Decision-Making in Large Language Models."
   NeurIPS 2025 GenAI4Health Workshop.

9. **G. Wan**, J. Allen, W. Ge, et al.
   "Two-step LightGBM for Human West Nile Virus Risk in Chicago."
   PLOS ONE 19(1): e0296283, 2024.

10. J.A. McCoy, L.D. Levine, **G. Wan**, et al.
    "Deep Learning on Intrapartum FHR to Predict Acidemia at Birth."
    American Journal of Obstetrics and Gynecology, 2024.

Honors & Awards

- Computational Health Informatics Fellowship, Harvard University (2022)
- Summa Cum Laude, University of Illinois at Urbana-Champaign (2020)
- SPIN Fellowship, National Center for Supercomputing Applications (2020)
- Best Presentation Nominee, Canadian Undergraduate Mathematics Conference (2020)
- Edmund J. James Scholar, University of Illinois at Urbana-Champaign (2018–2020)

Academic Service

Conference Reviewer: NeurIPS (2024, 2025), ICLR 2025, ICML 2025, AISTATS 2025
Journal Reviewer: IEEE Computational Intelligence Magazine
Teaching Assistant: UVA (CS 5012, DS 6310, DS 6030), Harvard (CS 109B)