

Contact Information

Salvatori Computer Science Center
Los Angeles, CA, 90089

✉ ollieliu@salvatori.com
✉ me@ollieliu.com

Research Interests

I'm broadly interested in multimodal foundation models. These days, I'm exploring their potential as agents of complex reasoning and scientific discovery. I'm particularly excited about::

- Designing and understanding architectures and algorithms that are applicable to scientific modalities, such as (meta)genomics, protein, physics, chemistry, and material sciences.
- Developing post-training and inference-time methods that enable FMs to solve complex reasoning and decision making problems.

Education

University of Southern California 08/2022 - Present

Ph.D. in Computer Science

Advisor(s): Dani Yogatama, Willie Neiswanger

Northwestern University

09/2019 - 06/2021

Master of Science in Industrial Engineering and Management Sciences

Advisor(s): Jorge Nocedal

Carnegie Mellon University

08/2013 - 08/2019

Master of Science in Machine Learning

Bachelor of Science in Statistics and Mathematics (with Honors)

Preprints

*: equal contribution; α - β : equal contribution, alphabetical order

Ollie Liu*, Deqing Fu*, Dani Yogatama, Willie Neiswanger “DeLLMa: Decision Making Under Uncertainty with Large Language Models” In: *arXiv Preprint*. [[📄 pdf](#)] [[🌐 website](#)]

Ollie Liu, Sami Jaghouar, Johannes Hagemann, Jeff Kaufman, Willie Neiswanger “MGFM: Metagenomic Foundation Model for Pandemic Monitoring” In: *Foundation Models for Science Workshop at NeurIPS 2024*. [[📄 pdf](#)]

Jiarui Zhang, **Ollie Liu**, Tianyu Yu, Jinyi Hu, Willie Neiswanger “Euclid: Supercharging Multimodal LLMs with Synthetic High-Fidelity Visual Descriptions” In: *arXiv Preprint*. [[📄 pdf](#)] [[🌐 website](#)] [[😊 models & datasets](#)]



Wenyue Hua, **Ollie Liu**, Lingyao Li, Alfonso Amayuelas, Julie Chen, Lucas Jiang, Mingyu Jin, Lizhou Fan, Fei Sun, William Yang Wang, Xintong Wang, Yongfeng Zhang “Game-Theoretic LLM: Agent Workflow for Negotiation Games” In: *arXiv Preprint*. [[📄 pdf](#)]



Ghazal Khalighinejad, Sharon Scott, **Ollie Liu**, Kelly L. Anderson, Rickard Stureborg, Aman Tyagi, Bhuwan Dhingra “MATVIX: Multimodal Information Extraction from Visually Rich Articles” In: *arXiv Preprint*. [[📄 pdf](#)] [[🌐 website](#)]



Publications


Deqing Fu $^{\alpha-\beta}$, Ruohao Guo $^{\alpha-\beta}$, Ghazal Khalighinejad $^{\alpha-\beta}$, **Ollie Liu** $^{\alpha-\beta}$, Bhuwan Dhingra, Dani Yogatama, Robin Jia, Willie Neiswanger “IsoBench: Benchmarking Multimodal Foundation Models on Isomorphic Representations” In: *Proceedings of COLM 2024*. [[📄 pdf](#)] [[🌐 website](#)]

Ting-Rui Chiang, Xinyan Velocity Yu, Joshua Robinson, **Ollie Liu**, Isabelle Lee, Dani Yogatama “On Retrieval Augmentation and the Limitations of Language Model Training” In: *Proceedings of*

NAACL 2024 (*short*). [ pdf] [ code]

Xianghao Kong*, **Ollie Liu***, Han Li, Dani Yogatama, Greg Ver Steeg “Interpretable Diffusion via Information Decomposition” In: *Proceedings of ICLR 2024*. [ pdf] [ code]

Ghazal Khalighinejad , **Ollie Liu**, Sam Wiseman “Approximating CKY with Transformers” In: *Proceedings of Findings of EMNLP 2023*. [ pdf] [ code]

Michael Hanna, **Ollie Liu**, Alexandre Variengien “How Does GPT-2 Compute Greater-Than?: Interpreting Mathematical Abilities in a Pre-Trained Language Model” In: *Proceedings of NeurIPS 2023*. [ pdf]

In Progress	Multimodal Foundation Model for Physical Sciences , Core Contributor, with Polymathic AI Self-Improving LM with Textual Feedback , Lead, with Microsoft Research	
Positions	New York University & Polymathic AI , Research Scientist	Fall 2024
	Microsoft Research, AI Frontiers , Research intern	Summer 2024
	Reka AI , Research intern	Summer 2023
	Redwood Research , Research Resident	Winter 2023
	Meta Reality Labs, Assistant Technologies , Applied Research Intern	Summer 2022
	You.com , Software Engineer Intern	Spring 2022
	Amazon Web Services, GenAI Innovation Center , Applied Scientist Intern	2021
Honors and Awards	University Organizer Fellowship , Open Philanthropy, 2024 Provost’s Fellowship , University of Southern California, 2022 Data Science Fellowship , Northwestern University, 2019 Royal E. Cabell Fellowship , Northwestern University, 2019 Senior Leadership Recognition , Carnegie Mellon University, 2018	
Invited Talks	USC Information Science Institute NLG Seminar , DeLLMa: A Framework for Decision Making Under Uncertainty with Large Language Models, [ vi deo]	
Teaching Experiences	TA , Machine Learning (Master), University of Southern California, Fall 2023 Co-Instructor , Machine Learning (Ph.D Elective), Northwestern University, Fall 2021 Co-Instructor , Mathematical Statistics (Ph.D Core), Northwestern University, Fall 2020 TA , Introduction to Machine Learning (Master), Carnegie Mellon University, 2 Semesters. TA , Principles of Computing (Undergraduate), Carnegie Mellon University, 4 Semesters.	
Doctoral Courseworks	Machine Learning, Natural Language Processing, Computer Vision, Learning Theory, Scalable Learning Systems, Theoretical Optimization	
Services and Activities	Reviewer , ICML (2023, 2024); NeurIPS (2024); ICLR (2025); ACL Rolling Review (2024) President , AI Safety Group , University of Southern California Student organizer , Center for Optimization and Statistical Learning , Northwestern University	
Skills	Software : Python (JAX, PyTorch,  Hugging Face), Linux, R, \LaTeX Language : Chinese (<i>native</i>), English (<i>proficient</i> , GRE V169+Q168)	