

Research Interests	I work at the intersection of decision-making, generative models, and AI-for-science, with the goal of building modular, auditable foundation models that support reliable reasoning and discovery. My work includes designing multimodal foundation models with applications in astronomy [C9] and metagenomics [P3], as well as developing tool-augmented methods for decision-making with foundation models [C6] and studying their underlying mechanisms [C1, C2, C3, C5, P4].	
Education	University of Southern California Ph.D. in Computer Science Advisors: Dani Yogatama, Willie Neiswanger	08/2022 - 05/2026
	Northwestern University M.S. in Industrial Engineering and Management Sciences	09/2019 - 06/2021
	Carnegie Mellon University B.S. (Hons) and M.S. in Machine Learning	08/2013 - 08/2019
Positions	FAIR at Meta , Research Intern Collaborator(s): Brandon Amos, Arman Zharmagambetov, Xian Li Developing algorithms that elicit optimization thinking in large language models.	05/2025 - 10/2025
	The Simons Foundation & Polymathic AI , Research Scientist Collaborator(s): François Lanusse, Shirley Ho Developed AION [P4], the first series of multimodal FMs for observational astronomy, leading post-training efforts and contributing to pre-training and model architecture ablations.	09/2024 - 12/2024
	Microsoft Research , Research Intern Collaborator(s): Corby Rosset Worked with the AI Frontiers Team and designed a self-improving workflow for post-training large language models from textual feedback, with applications in mathematical reasoning.	05/2024 - 08/2024
	Reka AI , Research Intern Collaborator(s): Qi Liu, Dani Yogatama Implemented training algorithms for reward models, with applications in multimodal data.	05/2023 - 07/2023
	Redwood Research , Research Resident Collaborator(s): The REMIX Research Program Applied path patching to identify reusable subgraph of language models that implement elementary mathematical operations, with results accepted to NeurIPS 2023 [C1].	01/2023 - 05/2023
Refereed Publications	<p>★: equal contribution; orange: featured publication</p> <p>[C11] Ang Li★, Charles L. Wang★, Kaiyu Yue★, Zikui Cai★, Ollie Liu★, Deqing Fu★, Peng Guo★, Wang Bill Zhu★, Vatsal Sharan, Robin Jia, Willie Neiswanger, Furong Huang, Tom Goldstein, Micah Goldblum “Zebra-CoT: A Dataset for Interleaved Vision Language Reasoning” In: <i>arXiv Preprint</i>. [pdf] [models & datasets]</p> <p>[C10] Shangshang Wang, Julian Asilis, Ömer Faruk Akgül, Enes Burak Bilgin, Ollie Liu, Willie Neiswanger “Tina: Tiny Reasoning Models via LoRA” In: <i>arXiv Preprint</i>. [pdf] [website]</p>	

[C9] The Polymathic AI Collaboration, François Lanusse*, Liam Holden Parker*, Jeff Shen*, **Ollie Liu**, Tom Hehir, Leopoldo Sarra, Lucas Thibaut Meyer, Micah Bowles, Sebastian Wagner-Carena, Helen Qu, Siavash Golkar, Alberto Bietti, Hatim Bourfoune, Pierre Cornette, Keiya Hirashima, Geraud Krawezik, Ruben Ohana, Nicholas Lourie, Michael McCabe, Rudy Morel, Payel Mukhopadhyay, Mariel Pettee, Bruno Régaldo-Saint Blancard, Kyunghyun Cho, Miles Cranmer, Shirley Ho “AION-1: Omnimodal Foundation Model for Astronomy” In: *Proceedings of NeurIPS 2025, Oral Presentation at the NeurIPS 2025 AI for Science Workshop.* [pdf] [website]

[C8] Xiaoyuan Zhu, Muru Zhang, **Ollie Liu**, Robin Jia, Willie Neiswanger “LLM Unlearning Without Expert Curated Dataset” In: *Proceedings of COLM 2025.* [pdf] [website]

[C7] Ghazal Khalighinejad, Sharon Scott, **Ollie Liu**, Kelly L. Anderson, Rickard Stureborg, Aman Tyagi, Bhuwan Dhingra “MATVix: Multimodal Information Extraction from Visually Rich Articles” In: *Proceedings of NAACL 2025.* [pdf] [website]

[C6] **Ollie Liu***, Deqing Fu*, Dani Yogatama, Willie Neiswanger “DeLLMa: Decision Making Under Uncertainty with Large Language Models” In: *Proceedings of ICLR 2025, Spotlight Presentation.* [pdf] [website]

[C5] Deqing Fu*, Ruohao Guo*, Ghazal Khalighinejad*, **Ollie Liu***, Bhuwan Dhingra, Dani Yogatama, Robin Jia, Willie Neiswanger “IsoBench: Benchmarking Multimodal Foundation Models on Isomorphic Representations” In: *Proceedings of COLM 2024.* [pdf] [website]

[C4] 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]

[C3] Xianghao Kong*, **Ollie Liu***, Han Li, Dani Yogatama, Greg Ver Steeg “Interpretable Diffusion via Information Decomposition” In: *Proceedings of ICLR 2024.* [pdf] [code]

[C2] Ghazal Khalighinejad , **Ollie Liu**, Sam Wiseman “Approximating CKY with Transformers” In: *Proceedings of Findings of EMNLP 2023.* [pdf] [code]

[C1] 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]

Preprints

[P11] The Polymathic AI Collaboration, Jeff Shen, François Lanusse, Liam Holden Parker, **Ollie Liu**, Tom Hehir, Leopoldo Sarra, Lucas Thibaut Meyer, Micah Bowles, Sebastian Wagner-Carena, Helen Qu, Siavash Golkar, Alberto Bietti, Hatim Bourfoune, Pierre Cornette, Keiya Hirashima, Geraud Krawezik, Ruben Ohana, Nicholas Lourie, Michael McCabe, Rudy Morel, Payel Mukhopadhyay, Mariel Pettee, Bruno Régaldo-Saint Blancard, Kyunghyun Cho, Miles Cranmer, Shirley Ho “Universal Spectral Tokenization via Self-Supervised Panchromatic Representation Learning” In: *arXiv Preprint & NeurIPS 2025 AI for Science Workshop.* [pdf]

- [P8] Shangshang Wang, **Ollie Liu**, Jiarui Zhang, Willie Neiswanger “OpenMETAGENE: Large-Scale, Diverse, and Open Data Recipes for Multimodal Metagenomics Models” In: *arXiv Preprint & NeurIPS 2025 AI for Science Workshop*. [[pdf](#)]
- [P7] **Ollie Liu**, Arman Zharmagambetov, Willie Neiswanger, Xian Li, Brandon Amos “OPTTHINKER: Optimization Reasoning with Large Language Models” In: *arXiv Preprint*. [[pdf](#)] [[website](#)]
- [P6] Bang Liu, Xinfeng Li, Jiayi Zhang, Jinlin Wang, Tanjin He, Sirui Hong, Hongzhang Liu, Shaokun Zhang, Kaitao Song, Kunlun Zhu, Yuheng Cheng, Suyuchen Wang, Xiaoqiang Wang, Yuyu Luo, Haibo Jin, Peiyan Zhang, **Ollie Liu**, Jiaqi Chen, Huan Zhang, Zhaoyang Yu, Haochen Shi, Boyan Li, Dekun Wu, Fengwei Teng, Xiaojun Jia, Jiawei Xu, Jinyu Xiang, Yizhang Lin, Tianming Liu, Tongliang Liu, Yu Su, Huan Sun, Glen Berseth, Jianyun Nie, Ian Foster, Logan Ward, Qingyun Wu, Yu Gu, Mingchen Zhuge, Xiangru Tang, Haohan Wang, Jiaxuan You, Chi Wang, Jian Pei, Qiang Yang, Xiaoliang Qi, Chenglin Wu “Advances and Challenges in Foundation Agents: From Brain-Inspired Intelligence to Evolutionary, Collaborative, and Safe Systems” In: *arXiv Preprint*. [[pdf](#)] [[website](#)]
- [P5] Shangshang Wang, Julian Asilis, Ömer Faruk Akgül, Enes Burak Bilgin, **Ollie Liu**, Deqing Fu, Willie Neiswanger “Resa: Transparent Reasoning Models via SAEs” In: *arXiv Preprint & NeurIPS 2025 Efficient Reasoning Workshop*. [[pdf](#)] [[website](#)] [[models & datasets](#)]
- [P4] Haosheng Gan*, Deqing Fu*, Julian Asilis*, **Ollie Liu***, Dani Yogatama, Vastal Sharan, Robin Jia, Willie Neiswanger “Textual Representations Alone Can Steer Multimodal Models” In: *arXiv Preprint*. [[pdf](#)] [[website](#)]
- [P3] **Ollie Liu**, Sami Jaghouar, Johannes Hagemann, Shangshang Wang, Jason Wiemels, Jeff Kaufman, Willie Neiswanger “METAGENE-1: Metagenomic Foundation Model for Pandemic Monitoring” In: *arXiv Preprint & NeurIPS 2024 Foundation Models for Science Workshop*. [[pdf](#)] [[website](#)] [[models & datasets](#)]
- [P2] 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](#)]
- [P1] 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](#)]
- Honors and Awards
Spotlight Presentation, DeLLMa: Decision Making Under Uncertainty with Large Language Models, The Thirteenth International Conference on Learning Representations (ICLR), 2025
Technical Innovation Fellowship, University of Southern California, 2025
Nebius Research Credits Program, est. \$200,000, Nebius Group, 2025
University Organizer Fellowship, est. \$30,000, Open Philanthropy, 2024
Provost’s Fellowship, est. \$2,000 per year, University of Southern California, 2022
Data Science Fellowship, est. \$10,000, Northwestern University, 2019
Royal E. Cabell Fellowship, est. \$10,000, Northwestern University, 2019
Senior Leadership Recognition, Carnegie Mellon University, 2018

Invited Talks

- CSCI 699: Probabilistic and Generative Models**, Guest Lecture on Flow Matching.
- Intel EAI Tech Talk**, DeLLMa: Decision Making Under Uncertainty with Large Language Models.
- USC Information Science Institute NLG Seminar**, DeLLMa: Decision Making Under Uncertainty with Large Language Models. [[video](#)]
- USC NLP Lunch**, IsoBench: Benchmarking Multimodal Foundation Models on Isomorphic Representations.
- USC NLP Lunch**, How Does GPT-2 Compute Greater-Than?: Interpreting Mathematical Abilities in a Pre-Trained Language Model.

Teaching Experiences

- TA**, Generative Models (Ph.D. Elective), University of Southern California, Spring 2025
- 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 (2024, 2025); NeurIPS (2024, 2025); ICLR (2025, 2026); ACL Rolling Review (2024, 2025)
- 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, L^AT_EX
- Language:** Chinese (*native*), English (*proficient*, GRE V169+Q168)