Xingyu Fu

Contact

Princeton AI Lab
41 William St
Princeton, NJ 08540

E-mail: fxylyly@gmail.com
Twitter: @XingyuFu2
Home: Webpage

Education

University of Pennsylvania, Philadelphia 2020 - 2025
Ph.D. candidate in Computer Science
Advisor: Dan Roth, Thesis: Bridging Perception and Reasoning in Multimodal Models

University of Illinois at Urbana-Champaign, Champaign 2017 - 2020
B.S. in Statistics & Computer Science, GPA: 3.95/4.0
Advisor: Prof. Jiawei Han

Employment

Princeton University | Princeton, NJ Jul 2025 - Present
Postdoc Fellow @ PLI (Advisors: Zhuang Liu, Danqi Chen, and Sanjeev Arora)

Microsoft Research | Redmond, WA May 2024 - Aug 2024
Research Intern (Mentor: Cha Zhang)

Amazon AWS AI Labs | New York City, NY May 2022 - Dec 2022
Applied Scientist Intern (Mentor: Zhiguo Wang)

University of Pennsylvania | Philadelphia, PA May 2019 - Aug 2019
Summer Research Intern (Advisor: Dan Roth)

University of Illinois at Urbana-Champaign | Champaign, IL May 2018 - 2019
Research Assistant (Advisors: Jiawei Han, Jingbo Shang)

Publications

Google Scholar https://scholar.google.com/citations?hl=en&user=5p_uBNQAAAAJ
Semantic Scholar https://www.semanticscholar.org/author/Xingyu-Fu/2078360

* denotes equal contribution

Preprints

[4] Bangzheng Li, Chen Qu, Jianmo Ni, Ian Miao, Liu Yang, **Xingyu Fu**, Muhao Chen, Derek Zhiyuan Cheng, Cheng Zhong. "Reinforced Attention Learning." *Submitted 2026.* [PDF]

[3] Bo Li, Yida Yin, Wenhao Chai, **Xingyu Fu***, Zhuang Liu*. "UEval: A Benchmark for Unified Multimodal Generation ." *Submitted 2026.* [PDF]

[2] Yida Yin, Harish Krishnakumar, Chung Peng Lee, Boya Zeng, Wenhao Chai, Shengbang Tong, Wenhu Chen, Hu Xu, **Xingyu Fu**, Gabriel Herbert Sarch, Aleksandra Korolova, Zhuang Liu. "WorldBench: A Challenging and Visually Diverse Multimodal Benchmark." *Submitted 2026.* [PDF]

[1] **Xingyu Fu**, Siyi Liu, Yinuo Xu, Pan Lu, Guangqiuse Hu, Tianbo Yang, Taran Anantasagar, Christopher Shen, Yikai Mao, Yuanzhe Liu, Keyush Shah, Chung Un Lee, Yejin Choi, James Zou, Dan Roth, Chris Callison-Burch. "Learning Human-Perceived Fakeness in AI-Generated Videos via Multimodal LLMs." *Submitted 2025.* [PDF]

Peer-Reviewed Publications

[14] Xuan Qi, Luxi He, Dan Roth, **Xingyu Fu**. "DataProphet: Demystifying Supervision Data Generalization in Multimodal LLMs." *ICLR 2026.* [PDF]

[13] **Xingyu Fu**, Minqian Liu, Zhengyuan Yang, John Corring, Lijuan Lu, Jianwei Yang, Dan Roth, Dinei Florencio, and Cha Zhang. "ReFocus: Visual Editing as a Chain of Thought for Structublack Image Understanding." *ICML 2025.* [PDF]

[12] Jialuo Li, Wenhao Chai, **Xingyu Fu**, Haiyang Xu, Saining Xie. "ScienceT2I: Addressing Scientific Illusions in Image Synthesis." *CVPR 2025.* [PDF]

[11] Fei Wang*, **Xingyu Fu***, James Y. Huang, Zekun Li, Qin Liu, Xiaogeng Liu, Mingyu Derek Ma, Nan Xu, Wenxuan Zhou, Kai Zhang, Tianyi Lorena Yan, Wenjie Jacky Mo, Hsiang-Hui Liu, Pan Lu, Chunyuan Li, Chaowei Xiao, Kai-Wei Chang, Dan Roth, Sheng Zhang, Hoifung Poon, and Muhao Chen. "MUIRBENCH: A Comprehensive Benchmark for Robust Multi-image Understanding." *ICLR 2025.* [PDF]

[10] Yushi Hu*, Weijia Shi*, **Xingyu Fu**, Dan Roth, Mari Ostendorf, Luke Zettlemoyer, Noah A Smith, and Ranjay Krishna. "Visual Sketchpad: Sketching as a Visual Chain of Thought for Multimodal Language Models." *NeurIPS 2024, Behavioral ML Workshop at NeurIPS 2024* [PDF]
> 250 Github Stars

[9] **Xingyu Fu**, Muyu He, Yujie Lu, William Yang Wang, and Dan Roth. "Commonsense-T2I Challenge: Can Text-to-Image Generation Models Understand Commonsense?." *COLM 2024.* [PDF]

[8] **Xingyu Fu***, Yushi Hu*, Bangzheng Li, Yu Feng, Haoyu Wang, Xudong Lin, Dan Roth, Noah A. Smith, Wei-Chiu Ma, and Ranjay Krishna. "BLINK: Multimodal Large Language Models Can See but Not Perceive." *ECCV 2024.* [PDF]
Spotlight presentation at Workshop cVinW@CVPR 2024
Spotlight paper at Harmonious
> 40K Total downloads on HuggingFace
HuggingFace Paper of the day

[7] Bangzheng Li, Ben Zhou, Fei Wang, **Xingyu Fu** Dan Roth, and Muhao Chen. "Deceptive Semantic Shortcuts on Reasoning Chains: How Far Can Models Go without Hallucination?." *NAACL 2024.* [PDF]

[6] Max Ku, Tianle Li, Kai Zhang, Yujie Lu, **Xingyu Fu**, Wenwen Zhuang, and Wenhu Chen. "ImagenHub: Standardizing the evaluation of conditional image generation models." *ICLR 2024.* [PDF]

[5] **Xingyu Fu**, Sheng Zhang, Gukyeong Kwon, Pramuditha Perera, Henghui Zhu, Yuhao Zhang, Alexander Hanbo Li, William Yang Wang, Zhiguo Wang, Vittorio Castelli, Patrick Ng, Dan Roth, and Bing Xiang. "Generate then Select: Open-ended Visual Question Answering Guided by World Knowledge." *in Proceedings of the Annual Meeting of the Association for Computational Linguistics, ACL 2023, Findings.* [PDF]

[4] **Xingyu Fu**, Ben Zhou, Sihao Chen, Mark Yatskar, and Dan Roth. "Dynamic Clue Bottlenecks: Towards Interpretable-by-Design Visual Question Answering." *Arxiv 2023.* [PDF]

[3] **Xingyu Fu**, Ben Zhou*, Ishaan Preetam Chandratreya*, Carl Vondrick, and Dan Roth. "There's a Time and Place for Reasoning Beyond the Image.." *in Proceedings of the Annual Meeting of the Association for Computational Linguistics, ACL 2022.* [PDF]
Oral Presentation

[2] **Xingyu Fu***, Weijia Shi*, Xiaodong Yu, Zian Zhao, Dan Roth. "Design Challenges in Low-resource Cross-lingual Entity Linking.." *in Proceedings of the Conference on Empirical Methods in Natural Language Processing, EMNLP 2020.* [PDF]

[1] Ahmed El-Kishky*, **Xingyu Fu***, Aseel Addawoody, Nahil Sobhy, Clare Vossz, and Jiawei Han. "Constrained Sequence-to-sequence Semitic Root Extraction for Enriching Word Embeddings.." *in Proceedings of the Fourth Arabic Natural Language Processing Workshop, WANLP@ACL 2019.* [PDF]

Invited Talks

Princeton Visual AI Lab, (Host: Olga Russakovsky) 06/2025
Title: *Bridging Perception and Reasoning in Multimodal Models*

University of Washington, (Host: Ranjay Krishna) 01/2025
Title: *Multimodal Understanding with World Knowledge*

University of Pennsylvania, (Host: Lingjie Liu) 01/2025
Title: *Multimodal Understanding with World Knowledge*

University of Pennsylvania, (Host: Kostas Daniilidis) 11/2024
Title: *Multimodal Understanding with World Knowledge*

University of Pennsylvania, CLunch 09/2024
Title: *Better Evaluations for Generative Multimodal Models.*

Microsoft Azure AI, AI reading group, (Host: Yijuan Lu) 06/2024
Title: *BLINK: Multimodal Large Language Models Can See but Not Perceive.*

Amazon AWS Responsible AI Group, (Host: Mathew Monfort) 07/2023
Title: *Generate then Select: Open-ended VQA Guided by World Knowledge.*

Professional Services

Reviewer/Program Committee

- ICML (2025)
- NeurIPS (2025)
- ICLR (2025)
- CVPR (2024, 2025)
- ICCV (2025)
- ACL (2021, 2022, 2023, 2024)
- EMNLP (2021, 2022, 2023, 2024)
- NAACL (2021, 2022, 2023, 2024)

Mentorship

Xuan Qi (2025), THU Yao Class BS.

Bo Li (2025), Fudan BS.

Yinuo Xu (2024), UPenn MS.

Yikai Mao (2024), UPenn MS. → LinkedIn ML Engineer.

Jialuo Li (2024), THU BS. → PhD at Georgia Institute of Technology.

Muyu He (2023-2024), UPenn MS. → Tiktok ML Engineer.

Ishaan Preetam Chandratreya (2023-2024), Columnbia BS. → PhD at MIT.

Teaching

Guest Lecture (*Instructor: Zhuang Liu*) — Fall 2025
COS 597K: Frontiers in Deep Learning, Princeton University

Teaching Assistant (*Instructor: Christopher Callison Burch*) — Fall 2024
CIS 8000: Research Practicum, UPenn

Teaching Assistant (*Instructor: Dan Roth*) — Spring 2023
CIS 7000: Machine Reasoning, UPenn

Teaching Assistant (*Instructor: Christopher Callison Burch*) — Fall 2021
CIS 5210: Artificial Intelligence, UPenn