

Gang Li

leemagpie@gmail.com | +1 302-367-6681 | linkedin | google scholar | website

EXPERIENCE

Research Scientist

Sep. 2024 – Present

Orby AI (acquired by Uniphore)

Mountain View, CA

- Distill from strong teacher models to open source models using browser-use trajectories. Obtained previous SoTA performance among open source models on *WebArena*, a complex Web agent benchmark
- Learn from environment feedback with Reinforcement Learning with Verifiable Rewards (RLVR) for multi-step tasks. Improved model performance on real-world tasks with RLVR using synthetic data (paper release soon)
- Lead development of Orby's post-training infra, which supports multi-node SFT and offline/online RLVR (PPO / GRPO). Stable training with 3B-72B multimodal models and Ray-based browser environments

Senior Software Engineer

Jan. 2020 – Sep. 2024

Google Research / DeepMind

Mountain View, CA

- Created the first *foundational UI understanding model*, with SoTA performance on various UI tasks
- Co-created the first *unified user modeling model* on diverse visual content, with SoTA performance on score, gaze and scanpath prediction
- Co-created the first *dataset and model* for rich human feedback on generative images, enabling accurate fix of problematic region without changing the entire image
- Launched automation models in Google products including Android Voice Access and UI-driven Pixel Testing; Used rich human feedback to improve Google Search and Google's text-to-image models
- Mentored the team on high-quality codebase for fast iterations, model finetuning for diverse product requirements, and stable/scalable model serving

Software Engineer

March 2018 – Jan 2020

Google Health

Mountain View, CA

- Created audio de-identification models and dataset for doctor-patient conversation
- Worked on entity and relation extraction models for doctor-patient conversation understanding

TECHNICAL SKILLS

Modeling: LLM pre-training and post-training, multi-turn RL, GUI agents, entity recognition, relation extraction

Data: large-scale high-quality pretraining data collection via crawling and synthetic data, LLM-based GUI interactions

Frameworks: PyTorch, Jax, Ray, Apache Beam, Amazon MTurk

Languages: Python, C++, Javascript, HTML/CSS, Java, SQL

EDUCATION

University of Delaware

Newark, DE

Ph.D. in Computer Science

Aug. 2011 – March 2018

Beijing Institute of Technology

Beijing, China

Bachelor of Engineering in Information Engineering

Aug. 2007 – May 2011

SELECTED PUBLICATIONS

- [1] **Gang Li** and Yang Li. Spotlight: Mobile UI understanding using vision-language models with a focus. In *The Eleventh International Conference on Learning Representations*, 2023. [Paper] [Google Research Blog].
- [2] Youwei Liang*, Junfeng He*, **Gang Li***, Peizhao Li, Arseniy Klimovskiy, Nicholas Carolan, Jiao Sun, Jordi Pont-Tuset, Sarah Young, Feng Yang, Junjie Ke, Krishnamurthy Dj Dvijotham, Katie Collins, Yiwen Luo, Yang Li, Kai J Kohlhoff, Deepak Ramachandran, and Vidhya Navalpakkam. Rich human feedback for text-to-image generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024. * **Co-first authors. Best Paper Award.** [Paper][Google Research Blog].