Xinyi Wang

xinyi_wang@ucsb.edu
Homepage: https://wangxinyilinda.github.io/
Tel.: (226)387-9013 or (805)455-1121

About Me

I am a fourth-year Ph.D. candidate in the computer science department at the University of California, Santa Barbara (UCSB). My research interest lies in developing principled understandings of foundation models to improve and make better use of them.

Education

- **University of California, Santa Barbara (UCSB)** — Santa Barbara, CA, US
  Ph.D. in Computer Science (expected) — *9.2020 - present*
  - GPA: 4.0/4.0
  - Advisor: William Yang Wang

- **The Hong Kong University of Science and Technology (HKUST)** — Hong Kong, China
  B.Sc in Applied Mathematics and Computer Science — *9.2016 - 7.2020*
  - GPA: 3.7/4.3

- **University of California, Los Angeles (UCLA)** — Los Angeles, CA, US
  Term exchange in Mathematics (Non-degree) — *9.2019 - 12.2019*
  - GPA: 3.9/4.0 (Dean's Honors List)

Honors and Awards

- J.P. Morgan AI PhD Fellowship — J.P. Morgan AI, 2024
- Computer Science Outstanding Publication Award — UCSB, 2024
- Second place in the Alexa Prize Simbot Challenge — Amazon, 2023
- Academic Excellence Fellowship — UCSB, 2020-2024
- Joseph Needham Merit Scholarship — Hong Kong, 2020-2024
- The 15th Epsilon Fund Award — HKUST, 2020
- Chern Class Scholarship — HKUST, 2017-2020
- HKSAR Government Scholarship Fund - Reaching Out Award — Hong Kong, 2019-2020
- University's Scholarship Scheme for Continuing Undergraduate Students — HKUST, 2017-2020

Publications

- Alon Albalak, Yanai Elazar, Sang Michael Xie, Shayne Longpre, Nathan Lambert, **Xinyi Wang**, Niklas Muennighoff, Bairu Hou, Liangming Pan, Haewon Jeong, Colin Raffel, Shiyu Chang, Tatsunori Hashimoto, William Yang Wang. *A Survey on Data Selection for Language Models.* TMLR 2024. [paper]

- **Xinyi Wang**, Alfonso Amayuelas, Kexun Zhang, Liangming Pan, Wenhu Chen, William Yang Wang. *Understanding the Reasoning Ability of Language Models From the Perspective of Reasoning Paths Aggregation.* ICML 2024. [paper]

- Iain Xie Weissburg, Mehir Arora, **Xinyi Wang**, Liangming Pan, William Yang Wang. *Tweets to Citations: Unveiling the Impact of Social Media Influencers on AI Research Visibility.* ICML 2024 (position paper). [paper]

- **Xinyi Wang**, Lucas Caccia, Oleksiy Ostapenko, Xingdi Yuan, William Yang Wang, Alessandro Sordoni. *Guiding Language Model Math Reasoning with Planning Tokens.* COLM 2024. [paper]

- Liangming Pan, Michael Saxon, Wenda Xu, Deepak Nathani, **Xinyi Wang**, William Yang Wang. *Automatically Correcting Large Language Models: Surveying the landscape of diverse self-correction strategies.* TACL 2023. [paper]

- Liangming Pan, Alon Albalak, **Xinyi Wang**, William Yang Wang. *Logic-LM: Empowering Large Language Models with Symbolic Solvers for Faithful Logical Reasoning.* Findings of EMNLP 2023. [paper]

- Wenhu Chen, Ming Yin, Max Ku, Pan Lu, Yixin Wan, Xueguang Ma, Jianyu Xu, **Xinyi Wang**, Tony Xia. *TheoremQA: A Theorem-driven Question Answering dataset.* EMNLP 2023. [paper]

- Wanrong Zhu, **Xinyi Wang**, Yujie Lu, Tsu-Jui Fu, Xin Eric Wang, Miguel Eckstein, William Yang Wang. *Collaborative Generative AI: Integrating GPT-k for Efficient Editing in Text-to-Image Generation.* EMNLP 2023. [paper]

- **Xinyi Wang**, Wanrong Zhu, William Wang. *Large Language Models Are Latent Variable Models: Explaining and Finding Good Demonstrations for In-Context Learning.* NeurIPS 2023, poster. [paper]

- Wenhu Chen, Xueguang Ma, **Xinyi Wang**, William W. Cohen. *Program of Thoughts Prompting: Disentangling Computation from Reasoning for Numerical Reasoning Tasks.* TMLR 2023. [paper]

- **Xinyi Wang**, Michael Saxon, Jiachen Li, Hongyang Zhang, Kun Zhang, William Yang Wang. *Causal Balancing for Domain Generalization.* ICLR 2023, poster. [paper]

- Michael Saxon, **Xinyi Wang**, Wenda Xu, William Yang Wang. *Relation Leakage in Elicited Natural Language Inference Datasets.* EACL 2023. [paper]

- Wenhu Chen, **Xinyi Wang**, William Yang Wang. *A Dataset for Answering Time-Sensitive Questions.* NeurIPS 2021 Datasets and Benchmarks Track, poster. [paper]

- **Xinyi Wang**, Wenhu Chen, Michael Saxon, William Yang Wang. *Counterfactual Maximum Likelihood Estimation for Training Deep Networks.* NeurIPS 2021, poster. [paper]

- Michael Saxon, Sharon Levy, **Xinyi Wang**, Alon Albalak, William Yang Wang. *Modeling Discolsive Transparency in NLP Application Descriptions.* EMNLP 2021, oral. [paper]

- **Xinyi Wang***, Haiqin Yang*, Liang Zhao, Yang Mo and Jianping Shen. *RefBERT: Compressing BERT by Referencing to Pre-computed Representations.* IJCNN 2021, oral. [1] [paper]

- **Xinyi Wang**, Yi Yang. *Neural Topic Model with Attention for Supervised Learning.* AISTATS 2020, poster. [paper]

INTERNSHIPS

- **Research Intern at MIT-IBM Watson Lab** — Boston, MA, US
  Mentor: Yikang Shen, Rameswar Panda — *6.2024 - 9.2024*
  - Topic: large language model related.

- **Research Intern at Microsoft Research** — Montreal, QC, Canada
  Mentor: Alessandro Sordoni — *6.2023 - 10.2023*
  - Topic: parameter-efficient fine-tuning to improve math reasoning ability of large language models.

SERVICES

- Conference reviewer: AAAI (2022, 2023), NeurIPS (2021, 2023, 2024), ICLR (2024), ICML (2024), COLM (2024)

- Journal reviewer: TPAMI (2024)

---

[1] * denotes equal contribution.