Huihan Li

huihanlhh.github.io

huihanli831@gmail.com

(774)-290-6300

Education

University of Southern California — Los Angeles, CA
PhD in Computer Science, Natural Language Processing — *Sep 2022 - Present*
Honors: Amazon Fellowship, 2024-2025
Advisor: Xiang Ren

Princeton University — Princeton, NJ
M.S.E. in Computer Science — *Sep 2020 - May 2022*
Honors: Siebel Scholars, Class of 2022
Advisor: Danqi Chen

Wellesley College — Wellesley, MA
B.A in Computer Science / Cognitive and Linguistic Sciences — *Sep 2016 - May 2020*
Honors: Sigma Xi Scientific Research Honor Society; Durant Scholars *magna cum laude*

Publications

Diagnosing Memorization in Chain-of-Thought Reasoning, One Token at a Time [pdf] [code]
Huihan Li*, You Chen*, Siyuan Wang, Yixin He, Ninareh Mehrabi, Rahul Gupta, Xiang Ren
EMNLP 2025

Attributing Culture-Conditioned Generations to Pretraining Corpora [pdf] [code]
Huihan Li*, Arnav Goel*, Keyu He, Xiang Ren
ICLR 2025

In Search of the Long-Tail: Systematic Generation of Long-Tail Inferential Knowledge via Logical Rule Guided Search [pdf] [code]
Huihan Li, Yuting Ning, Zeyi Liao, Siyuan Wang, Xiang Lorraine Li, Ximing Lu, Wenting Zhao, Faeze Brahman, Yejin Choi, Xiang Ren
EMNLP 2024

CULTURE-GEN: Revealing Global Cultural Perception in Language Models through Natural Language Prompting [pdf] [code]
Huihan Li, Liwei Jiang, Jena D. Hwang, Hyunwoo Kim, Sebastin Santy, Taylor Sorensen, Bill Yuchen Lin, Nouha Dziri, Xiang Ren, Yejin Choi
COLM 2024

Ditch the Gold Standard: Re-evaluating Conversational Question Answering [pdf] [code]
Huihan Li*, Tianyu Gao*, Manan Goenka, Danqi Chen
ACL 2022
Outstanding Paper Award

Controllable Text Generation with Language Constraints [pdf]
Howard Chen, **Huihan Li**, Danqi Chen, Karthik Narasimhan
arxiv, 2022

Internship Experience

Meta, GenAI — Bellevue, WA
*Research Intern with **Liang Tan** and **Bo Xiong*** — *May 2025 – Aug 2025*
- Independently designed and implemented an improved reinforcement learning sampling method to enhance mathematical reasoning in large language models

Allen Institute of Artificial Intelligence, Mosaic — Seattle, WA
*Research Intern with **Yejin Choi*** — *May 2023 – Mar 2024*
- Designed, led and implemented research project on studying **multicultural bias** in large language models

Apple, AI/ML — Seattle, WA
AI/ML Intern, Answers and Web Ranking Team, AI/ML Information Intelligence — *May 2022 – Aug 2022*

- Collected annotation data for training **multilingual question answering systems**
- Implemented **multilingual LUKE** model in **Flax**
- Wrote trainer pipeline and trained multilingual extractive question answering model with **TensorFlow** and **JAX**

Google Pittsburgh, PA
Software Engineering Intern *May 2019 – Aug 2019*

- Designed, implemented and tested an **end-to-end dialog feature** for price queries on Google Assistant devices
- Created **natural language understanding grammar** that recognizes user intent to ask for price and extract key information from query

Seedlink Technology Shanghai, China
Linguistic Research Intern *Jun 2017 – Aug 2017*

- Researched on **extracting linguistic features** from written text that correlated with human personality
- Improved the **segmentation** accuracy of in-house Chinese dictionary model by 11% by creating more accurate user dictionary

Skills

- **Frameworks**: PyTorch, Jax, Flax, TensorFlow, CUDA, Pandas, NumPy
- **Annoation**: Amazon Mechanical Turk, Prolific
- **Languages**: Python, Java, C, C++, Assembly, Racket, Standard ML, HTML, CSS
- **DevOps**: Amazon Web Services, Microsoft Azure, Git

Services

- **Reviewing**: COLING 2024, ACL 2024, EMNLP 2024, COLING 2025, ICLR 2025, ARR August 2024, ARR December 2024, NAACL 2025 Demo Track, ACL 2025, COLM 2025, EMNLP 2025

Teaching

Princeton University, Computer Science Princeton, NJ
Teaching Assistant *Sep 2020 – May 2022*

- Led Discussion Sessions and graded assignments for Princeton's undergraduate course COS217, Introduction to Programming Systems.

Wellesley College, Computer Science Wellesley, MA
Tutor & Grader *Jan 2018 – Dec 2019*

- Led Office Hours and graded assignments for Wellesley's undergraduate course CS230, Data Structures, and CS235, Theory of Computation.