

HAIWEN LI

+1-805-637-5198 | haiwenli@mit.edu

Cambridge, MA 02139

RESEARCH INTERESTS

Human-AI/LLM Interaction, AI Safety, Social Computing

EDUCATION

- **Massachusetts Institute of Technology** September 2022 - June 2027 (expected)
Ph.D. in Social and Engineering Systems Cambridge, MA
 - GPA: 5.0/5.0
 - Relevant coursework: Probability Theory, Statistical Inference, Information Theory, Econometrics and Causal Inference
- **University of California, Los Angeles** June 2022
B.S. in Statistics (UCLA Honors Program) Los Angeles, CA
 - GPA: 3.96/4.0 (Summa Cum Laude)
 - Relevant coursework: Statistical Modeling and Inference, Statistical Models in Finance, Machine Learning, Reinforcement Learning

PUBLICATION

Author names: **bold** indicates first/lead author.

Scaling Human Judgment in Community Notes with LLMs [\[Paper\]](#)

Haiwen Li, Soham De, Manon Revel, Andreas Haupt, Brad Miller, Keith Coleman, Jay Baxter, Martin Saveski, Michiel Bakker
Journal of Online Trust and Safety (2025)

Nudge Users to Catch Generative AI Errors [\[Paper\]](#)

Renée Gosline, Yunhao Zhang, Haiwen Li, Accenture Research
MIT Sloan Management Review (2024)

WORK IN PROGRESS

Reinforcement Learning from Community Feedback

(with Michiel Bakker, Martin Saveski, Jay Baxter, and Soham De)

Improving Accuracy and Bridging: A Case Study of LLM Assistants for X Community Notes Ratings

(with Michiel Bakker, Martin Saveski, Jay Baxter, and Soham De)

Trusting GenAI: A Large Scale Generative Search Experiment

Haiwen Li, Sinan Aral
Under review at *Nature Human Behavior*

RESEARCH EXPERIENCE

- **Massachusetts Institute of Technology** Fall 2023 -
Research Assistant Cambridge, MA
- **Visual Intelligence Lab (UCLA)** September 2021 - March 2022
Research Assistant Los Angeles, CA
 - Implemented reinforcement learning algorithms and neural networks to construct a solver that significantly reduces the computational cost of large-scale partially observable MDP
- **Center for Vision, Cognition, Learning, and Autonomy (UCLA)** June 2021 - August 2021
Research Assistant for Reasoning in Vision, Language, and Knowledge Los Angeles, CA
 - Created interactive visualizations for a Question-Answering dataset
- **MIT Sloan School of Management** March 2020 - October 2020
Research Assistant Remote
 - Developed text-based product classifier using latent semantic analysis and XGBoost algorithm
 - Studied how tiktok video features drive video engagement and product sales

TEACHING EXPERIENCE

- MIT Sloan School of Management

Teaching Assistant for Digital Marketing and Social Media Analytics (Student Feedback: 6.3/7.0)

◦ Conducted recitation sessions on empirical causal inference analysis
- Fall 2024
Cambridge, MA

INVITED PRESENTATIONS

- The 2025 Trust and Safety Research Conference
- 2025

HONORS AND AWARDS

- Michael Hammer Fellowship

Institute for Data, Systems, and Society, MIT
- 2022
- Dean’s Honors List

University of California, Los Angeles
- 2020, 2021

SERVICE

- Reviewer
- CSCW 2025, HAI 2025
- 2025