# JIAYI WU

Phone: (+1) 401-743-3391 ⋄ Email: jiayi_wu4@brown.edu
Homepage: https://jiayiw005.github.io/
Google Scholar ⋄ GitHub ⋄ LinkedIn

## EDUCATION

**Brown University**                                    Sept 2024 - May 2028 (expected)
Sc.B. Mathematics-Computer Science and A.B. Applied Mathematics
Related Coursework:

- Machine Learning / Artificial Intelligence: CSCI 2952N Advanced Topics in Deep Learning, CSCI 2470 Deep Learning (Graduate Level), CSCI 2952W Critical Data and Machine Learning Studies

- Theory / Formal Methods: CSCI 1010 Theory of Computation, CSCI 1715 Formal Proof and Verification, CSCI 1710 Logic for Systems, CSCI 1951Y Using an Interactive Proof Assistant to Do Mathematics, CSCI 1973 Independent Study (on Automated Theorem Proving)

- Mathematics: MATH 1530 Abstract Algebra, APMA 1655 Introduction to Probability and Statistics with Theory, MATH 1030 Graph Theory, MATH 0540 Linear Algebra with Theory

**Shanghai Qibao Dwight High School**                              Sept 2021 - May 2024
International Baccalaureate Diploma Programme (IBDP), Full Diploma

## RESEARCH INTERESTS

I hope to build the formal and empirical infrastructure for **trustworthy machine learning** across two complementary dimensions:

- on the *learning* side, model architectures informed by structured latent representations and neurosymbolic approaches, and how they support more robust and interpretable inference and decision-making.

- on the *guarantee* side, evaluation and auditing pipeline design informed by formal verification, and how they scaffold more aligned, accountable automated systems.

## PUBLICATIONS

\* represents equal contribution.

[1] E. Poole-Dayan, **Jiayi Wu**, J. Pei, and M. A. Bakker, "Benchmarking overton pluralism in LLMs," in *NeurIPS 2025 Workshop on Evaluating the Evolving LLM Lifecycle: Benchmarks, Emergent Abilities, and Scaling*, 2025. [Online]. Available: https://arxiv.org/abs/2512.01351.

[2] C. J. Li\*, **Jiayi Wu\***, Z. Mo, A. Qu, Y. Tang, K. I. Zhao, Y. Gan, J. Fan, J. Yu, J. Zhao, P. P. Liang, L. A. A. Pastor, and K. Larson, "Simulating society requires simulating thought," in *The Thirty-Ninth Annual Conference on Neural Information Processing Systems Position Paper Track*, 2025. [Online]. Available: https://arxiv.org/abs/2506.06958.

[3] C. J. Li\*, Z. Mo\*, Y. Tang, A. Qu, **Jiayi Wu**, K. I. Zhao, Y. Gan, J. Fan, J. Yu, H. Jiang, P. P. Liang, J. Zhao, L. A. A. Pastor, and K. Larson, "HugAgent: Evaluating LLMs in simulating individual-level human reasoning on open-ended tasks," in *NeurIPS 2025 Workshop on Bridging Language, Agent, and World Models for Reasoning and Planning (Spotlight)*, also accepted to *NeurIPS 2025 Workshop on PersonaLLM: Workshop on LLM Persona Modeling (Oral)* and *NeurIPS 2025 Workshop on Socially Responsible and Trustworthy Foundation Models (ResponsibleFM)*, 2025. [Online]. Available: https://arxiv.org/abs/2510.15144.

## RESEARCH EXPERIENCE

**Pattern Mining and Automated Tactic Discovery in Theorem Proving**  Oct 2025 - Present
*Brown University*                                                        Providence, RI

Research project in collaboration with Gavin Zhao, advised by Prof. Robert Lewis and Prof. Stephen Bach at Brown University Department of Computer Science.

**Formalizing and Benchmarking Overton Pluralistic Alignment in Large Language Model Evaluation [1]**                                                May 2025 - Present
*Massachusetts Institute of Technology*                                  Cambridge, MA

Visiting summer researcher project at MIT Sloan School of Management, Institute for Data, Systems, and Society (IDSS), and Center for Constructive Communication (CCC) at Media Lab, advised by Prof. Michiel Bakker, in collaboration with Elinor Poole-Dayan *et al.*

**Motif of Thoughts (MoT): Reusable Abstractions for Compositional Neurosymbolic Reasoning [2] [3]**                                              Mar 2025 - Present
*Massachusetts Institute of Technology*                                  Cambridge, MA

Research project in collaboration with Chance Jiajie Li *et al.* at MIT Media Lab, Department of Electrical Engineering and Computer Science (EECS), and Department of Brain and Cognitive Sciences (BCS).

**Pluggable Analyses for Modern Real Systems**                          Dec 2025 - Present
*Brown University*                                                        Providence, RI

Undergraduate Teaching and Research Awards (UTRA) project advised by Prof. Nikos Vasilakis at ATLAS Group, Brown University Department of Computer Science.

**Algorithmic Fairness and AI Governance Primer, Brown CS Socially Responsible Computing (SRC) Handbook**                                     Dec 2024 - Present
*Brown University*                                                        Providence, RI

Research project advised by Prof. Suresh Venkatasubramanian and Prof. Julia Netter at Brown University Department of Computer Science and Center for Tech Responsibility (CNTR).

**EncyclopAIdia: AI Policy Guides and Glossary**                       April 2025 - Present
*Carnegie Mellon University*                                                    Remote

Research project advised by Prof. Sarah Cen at CMU Department of Electrical and Computer Engineering (ECE) and Engineering & Public Policy (EPP).

**Trade-Off Modeling and Auditing Formalization in Platform-Based Algorithmic Management: A Multi-Objective Optimization Approach**             Nov 2024 - May 2025
*Brown University*                                                        Providence, RI

Research project advised by Prof. Harini Suresh at the Data in Society Collective (DISCO Lab), Brown University Department of Computer Science; previously class final project at CSCI 2952W Critical Data and Machine Learning.

## PROJECT EXPERIENCE

**Probing Structural Signals in Lean4 Proof Graphs with GNNs**          Sept 2025 - Dec 2025
*Brown University*                                                        Providence, RI

Class final project at CSCI 2470 Deep Learning, Brown University Department of Computer Science.

**Formalizing Foundations of Projective Geometry in Isabelle**    Sept 2025 - Dec 2025
*Brown University*    Providence, RI

Class final project at CSCI 1951Y Using an Interactive Proof Assistant to Do Mathematics, Brown University Department of Computer Science.

**ConnectRI Broadband Navigator Accessibility Audit**    Feb 2025 - May 2025
*Brown Initiative for Policy*    Providence, RI

Policy initiative at Digital Equity Team.

**AsyLex Refugee Claim Entity Extraction & Judgement Prediction**    Sept 2024 - Mar 2025
*AI4ALL Ignite Fellowship*    Remote

## LEADERSHIP & TEACHING

**Teaching Assistant**, CSCI 1952B Responsible Computer Science in Practice    Dec 2025 - Present
**Co-President**, AI Robotics Ethics Society (AIRES) @ Brown    Sep 2025 - Present
**Teaching Assistant**, CSCI 0170 Computer Science: An Integrated Introduction    Sep 2025 - Dec 2025
**AI Governance Panel Director**, Brown China Summit    Sep 2024 - Mar 2025

## TECHNICAL SKILLS

| | |
|---|---|
| **Programming Languages** | Java, Python, JavaScript, HTML/CSS, SQL/MySQL, MATLAB, Lean, Isabelle, ReasonML, Racket, C/C++, Processing |
| **Programming Tools** | PyTorch, TensorFlow, scikit-learn, NumPy, pandas, D3.js, Plotly, Vue.js, Vite, Git/GitHub, Linux |