ZINI YANG

Email: zini.yang@duke.edu | Homepage: https://ziniyang.github.io/zini_yang/

RESEARCH INTEREST

My research interests are: (1) **Social Science for Trustworthy AI:** Utilize insights from game theory, behavioral economics, and network science to bridge the gaps between LLMs and human intelligence; (2) **AI for Science & Social Science:** Explore how AI can be used to enhance scientific discovery and social science, and improve the modeling and simulations of human behavior.

EDUCATION

Duke University — Durham, United States
M.S. in Computer Science and Economics — Aug 2024 – May 2026 (Expected)
Advisor: Dr. Emily Wenger
Major GPA: 3.96/4.00
Honors & Awards: Master's Scholarship Award, Dean's Research Award

Shanghai University of Finance and Economics — Shanghai, China
B.S. in Mathematical Economics — Sep 2019 – Jul 2023
Advisor: Dr. Simin He
Major GPA: 3.73/4.00, Rank: 2/40
Honors & Awards: Graduate with Distinction, Academic Performance Scholarship (2019-2023)

PUBLICATIONS AND MANUSCRIPTS

Zini Yang, Richard So, and Emily Wenger. *Creativity Coverage: Human-Grounded Boundaries for Evaluating LLM Creativity*. Under submission to ACL 2026.

Randy Davila, Jesús A. De Loera, Jillian Eddy, Ethan X. Fang, Junwei Lu, and **Zini Yang**. *Inequality Ranking and Inference System (IRIS): Giving Mathematical Conjectures Numerical Value*. ICML Math-AI Workshop, 2025. [paper]

RESEARCH EXPERIENCE

Research Area 1: Social Science for Trustworthy AI

LLMs Creativity Evaluation — Durham, NC
Advisor: Dr. Emily Wenger, Dr. Richard So — *Oct 2025 – Present*

- Designed an improved LLM-creativity evaluation method that integrates human responses.
- Introduced LLM coverage rate, a new metric that provides more meaningful creativity assessment than existing methods.
- Benchmarked creativity of 6 LLMs across various tasks and found that all models achieved coverage rates below 60%, with performance dropping to about 10% on open-ended creative-writing tasks.
- Exposed limitations in the creative capabilities of the tested LLMs, particularly for open-ended writing tasks.

Data Protection Mechnism Design — Durham, NC
Advisor: Dr. Emily Wenger, Dr. Curtis R. Taylor — *Feb 2025 – Present*

- Proposed a data-protection framework to provide better user security.
- Developed a dynamic game-theoretic framework to model protector-attacker interactions in a sequential decision process.
- Solved for Markov Perfect Equilibrium which revealed the overall long-run protection inefficiency resulted from free-riding incentives among heterogeneous protectors.
- Derived actionable insights that help users make more effective data-protection decisions to deter attacks.

Research Area 2: AI for Science & Social Science:

AI for Mathematical Conjecture Discovery and Counterexample Finding — Durham, NC
Advisors: Dr. Ethan Fang; Dr. Junwei Lu — *Oct 2024 – Present*

- Published a paper at *ICML 2025 Workshop.*
- Introduced a new conjecture-quality evaluation metric that provides a scalable conjecture-scoring tool for the `graffitiai` system.
- Developed an RL-based counterexample-searching agent that achieved ∼95% success rate in disproving tested conjectures.
- Built a self-play system to generate high-quality conjectures, where a conjecture generator and a counterexample searcher iteratively propose and test conjectures.

LLM Based Vaccine Warning System Durham, NC

Duke NLP Group, Advisor: Dr. Bhuwan Dhingra *Jul 2025 – Present*

- Built an AI-driven method to detect early vaccine-related concerns from large-scale web text.
- Designed a scalable two-stage pipeline to extract vaccine-relevant content from highly imbalanced Common Crawl data ($<0.1\%$ relevant), using fastText for initial filtering and Qwen for multi-category classification.
- Applied UMAP and HDBSCAN to cluster similar vaccine-related content, and generated trend reports through LLM summarization.
- Validated the performance of our method on monitoring vaccine safety under the guidance of medical professionals.

Research Area 3: Behavioral and Network Economics

Expectation Formation with Correlated Variables Shanghai, China

Research Intern, Advisor: Dr. Simin He *Jul 2022 – Oct 2023*

- Conducted an empirical analysis of how individuals form expectations about a certain variable using information from multiple, possibly correlated sources.
- Developed and implemented the experiment with oTree to collect human responses from treatment and control groups.
- Conducted group-level significance analysis using OLS regressions and Mann–Whitney tests to compare experimental treatments.
- Built an end-to-end pipeline (Bash/Makefile + R/renv) for cleaning, estimation, figure/report for review.

Public Goods Provision in a Network Formation Game Shanghai, China

Research Intern, Advisor: Dr. Simin He *May 2022 – Nov 2022*

- Analyzed how people form and retain costly, mutually-agreed social connections to cooperate on providing social goods.
- Preprocessed oTree logs, reconstructed bilateral link networks from individual responses, and conducted network-level statistical analysis.
- Discovered through analysis that participants favored simple equal-splitting the cost of public goods over the more efficient strategy of providing the good sequentially.

INDUSTRY EXPERIENCE

Xiaohongshu (rednote) Shanghai, China

Data Analyst Intern *May 2023 – Aug 2023*

- Implemented a consumer-behavior model for consumption-trend analysis in SQL/Python, used to improve 20+ online/offline acquisition campaigns.
- Improved the recommendation system in collaboration with the IT team.
- Improved the "Find a Partner" sub-platform, resulting in 0.4M+ additional user engagements.

TEACHING EXPERIENCE

COMPSCI 653: Human-Centered Computing Duke University

Teaching Assistant *Fall 2025*

COMPSCI 230: Discrete Mathematics for Computer Science Duke University

Teaching Assistant *Spring 2025*

SKILLS

Statistical & Analytical: GLM (logistic/Poisson/NegBin), hypothesis testing, causal inference, A/B testing, time-series analysis, dimensionality reduction & clustering (PCA, k-means/HDBSCAN)
Programming: Python (NumPy, pandas, SciPy, Matplotlib, Seaborn, scikit-learn, PyTorch), C++, SQL, R, SAS, Java
Machine Learning: Transformers & Fine-tuning (LoRA/QLoRA), LSTM/RNN, XGBoost/Random Forest/Decision Tree, RAG