

Jiaqi Lai

Singapore | jiaqi.lai.29@gmail.com

PROFILE

Researcher using formal and computational methods to study how AI-related disruptions interact with public trust, information environments, and policy implementation. Develops evidence-constrained approaches for evaluating AI-enabled systems in complex social settings, with a focus on robustness, uncertainty, and human oversight. Research experience spans computational social science, policy analysis, and human-centred AI in Singapore, China, and the UK.

SELECTED RESEARCH OUTPUTS

Lai, J., Hou, L., & Huang, W. (2026). “Stability of AI Governance Systems: A Coupled Dynamics Model of Public Trust and Social Disruptions.” [arXiv:2603.20248](https://arxiv.org/abs/2603.20248).

First-author work deriving conditions under which public-trust systems recover from, or depart persistently after, AI-related disruptions; combines formal analysis with reproducible simulations across network structures.

Steed, A., & **Lai, J.** (2025). “Comparison of hand tracking-based and controller-based interaction in a consumer virtual reality game.” *Virtual Reality*, 29, 120. [JCR Q1] [doi:10.1007/s10055-025-01190-5](https://doi.org/10.1007/s10055-025-01190-5).

RESEARCH & POLICY EXPERIENCE

AI Singapore

Research Assistant, under Dr. William Tjhi

Singapore

Apr. 2026 – Present

- Designed and implemented a human-governed extension layer for an existing policy-simulation ABM: mechanism-scope gaps trigger LLM-authored candidate modules within a bounded declarative interface.
- Built an auditable staged workflow with provenance labels, deterministic invariant checks, paired-seed causal-direction probes, multi-seed stress tests, versioned artifacts, and explicit human activation.
- Implemented sealed historical reconstruction and blind-forecast backtesting tracks to evaluate policy reasoning without treating similarity to historical policy as success.

Nanyang Technological University, School of Economics

Independent Research Project, PhD-level Mathematical Economics

Singapore

Jan. 2025 – Jul. 2025

- Developed a coupled model of public trust and AI-related events to examine when a governance system can recover from shocks and when feedback produces persistent departure.
- Derived closed-form equilibria, an exact spectral recoverability criterion ($\rho(J_{2n}) < 1$), and a conditional bound on topology sensitivity; tested self-excitation, memory, heterogeneous response, and network effects through reproducible simulations.
- Separated equilibrium heterogeneity from stability across random, echo-chamber, and star/KOL network structures.

China Development Research Foundation

Policy Research Assistant

Beijing, China

Oct. 2024 – Mar. 2025

- Evaluated gaps between AI-policy intent and deployment in education and rural contexts, identifying structural barriers and incentive misalignments.
- Synthesised cross-disciplinary evidence into policy evaluation reports for senior stakeholders, defining success indicators and risk frameworks for AI policy programmes.

Nanyang Technological University, WKCSC

Machine Learning Researcher

Singapore

Sep. 2024 – Apr. 2025

- Analysed cross-platform differences in AI-related sentiment, topics, and information diffusion on Weibo and X/Twitter, comparing public discourse across linguistic and information environments.
- Built cross-platform NLP pipelines and benchmarked transformer and recurrent classifiers for cross-lingual sentiment analysis under distribution shift.

RESEARCH & POLICY EXPERIENCE (CONTINUED)

UCL Interaction Centre (UCLIC)

Research Assistant, under Prof. Yvonne Rogers

London, UK

Sep. 2022 – Apr. 2024

- Designed and executed a mixed-methods within-subject experiment ($N = 86$) comparing LLM-powered and branching-dialogue agents, examining how perceived autonomy relates to trust and satisfaction.
- Applied non-parametric inference and moderated regression to autonomy–trust–satisfaction pathways, complemented by thematic analysis of open-ended responses.
- Found that greater perceived autonomy did not increase trust or satisfaction, highlighting the need for capability-sensitive evaluation and caution when measuring latent trust constructs.

SELECTED WRITING

Lai, J., Wang, T., & Li, X. (2024). “The Disinformation Campaigns about the Israel–Palestine Conflict in the Digital Era.” *MSc Critical Inquiry Project*, awarded full marks; available upon request.

Analysed the impact of generative AI on state-sponsored computational propaganda through Strategic Narrative Warfare theory.

ADDITIONAL EXPERIENCE

Golden Gate Ventures

Machine Learning Intern

Singapore

Jun. 2025 – Aug. 2025

- Developed a multi-agent research pipeline over heterogeneous data sources and fine-tuned BERT and Qwen for domain-specific classification.
- Gained production-facing experience of model selection, output alignment, data quality, and stakeholder evaluation in AI-assisted research workflows.

UCL Computer Science

Research Assistant

London, UK

Oct. 2023 – May 2024

- Analysed behavioural logs and prepared figures for the consumer-VR study resulting in the peer-reviewed publication listed above.

Canopy

Data & Sales Enablement Executive

Singapore

Aug. 2025 – Jan. 2026

- Developed a quantitative pricing model using logistic regression and generalised additive models; coordinated cross-functional data integration for investor due diligence.
- Worked with operational teams to translate model assumptions and outputs into decision-relevant evidence for internal and investor-facing use.

EDUCATION

Nanyang Technological University | Singapore

MSc Information Studies

Aug. 2024 – Jul. 2025

- Completed **PhD-level Mathematical Economics by special permission** with grade **A (5/5)**.
- Selected coursework: Machine Learning Implementation; Text & Web Mining; Disinformation & Information Professions (**A, 5/5**).

University College London | London, UK

BSc Psychology & Language Sciences

Sep. 2021 – Jun. 2024

- Selected coursework: Deep Learning & Language Processing (1st Class); R & Statistical Methods (1st Class); Human-Computer Interaction.

METHODS

AI Evaluation & Governance: Evaluator isolation; provenance and versioned lineage; deterministic, invariant, causal-direction, and multi-seed stress tests; historical backtesting; human-gated activation.

Formal & Empirical Methods: Agent-based and multi-agent simulation; discrete-time dynamical systems; network and spectral analysis; sensitivity analysis; experimental design; non-parametric inference; regression; thematic analysis.

Machine Learning & Programming: Python, R, SQL, Git; NLP and cross-lingual analysis; BERT/Qwen fine-tuning; reproducible simulation workflows.