

Curriculum Vitæ

Patrick Zhao

patrick.zhao@sfu.ca

see [my website](#) for a more casual viewing experience

research interests

human-centered computing and artificial intelligence; agent-based modeling and computational social science; data valuation and governance; AI safety and interpretability

education

Simon Fraser University

MSc. Thesis, Computer Science; Burnaby, BC, Canada; Jan 2025 - Current.

People and Data Computing Group

Simon Fraser University

BSc., Computer Science; Burnaby, BC, Canada; Sept 2019 - Aug 2024.

publications

Mechanism Plausibility in Generative Agent-Based Modeling

Patrick Zhao, David Huu Pham, Nicholas Vincent

ACM FAccT 2026

Also at PoliSim Workshop @ CHI 2026 (Best Paper)

Empowering Users Together: Connecting Algorithmic Collective Action and Explainable AI

Ayana Hussain, Cole Thacker, Patrick Zhao, Aditya Karan, Nicholas Vincent

Algorithmic Collective Action Workshop @ NeurIPS 2025

Let Them Down Easy! Contextual Effects of LLM Guardrails on User Perceptions and Preferences

Mingqian Zheng, Wenjia Hu, Patrick Zhao, Motahhare Eslami, Jena D. Hwang, Faeze Brahman, Carolyn Rose, Maarten Sap

EMNLP Findings 2025

An Audit and Analysis of LLM-Assisted Health Misinformation Jailbreaks Against LLMs

Ayana Hussain, Patrick Zhao, Nicholas Vincent

AIES 2025

research experience

Simon Fraser University, People and Data Computing Group

Graduate Research Student. Advisor: Nicholas Vincent

Burnaby, BC, Canada; Jan 2025 - Current

Human-centered AI, AI safety, and computational social science. Currently exploring the epistemological and methodological challenges that arise when using large language models as cognitive agents in simulations.

Carnegie Mellon University, Language Technologies Institute

Research Assistant.

Burnaby, BC, Canada; Sept 2024 - April 2025

Designed and engineered chatbot interfaces with safety mechanisms for experimental human-subject studies.

Simon Fraser University, Beedie School of Business

Research Assistant. Advisor: Peng Zhang

Burnaby, BC, Canada; Mar 2022 - Dec 2022

Fine-tuned early GPTs (text-davinci-003 and previous) to extract data from the Johannesburg stock exchange. Examined educational origins of executives leading publicly traded companies in South Africa.

professional experience

SAP

Software Developer Intern

Vancouver, BC, Canada

Jan 2023 - Aug 2023

Kongsberg Maritime

Software Developer and Test Intern

Port Coquitlam, BC, Canada

May 2021 - Aug 2021

additional training

Center for AI Safety

AI Safety, Ethics & Society Program (Spring '25 cohort)

Remote

Feb 2025 - May 2025

leadership and service

Co-founder, Executive Director

The Pollinator Project

Coquitlam, BC, Canada

Oct 2016 - Aug 2021

news/media coverage

- [Anthropic's Claude AI Can Now End Abusive Conversations For 'Model Welfare'](#)
Forbes. Victor Dey. Aug 21, 2025
- [Planet Superheroes](#)
Junior Scholastic. Jessica Press. April 22, 2019
- [Bee army sets out to transform Coquitlam gardens](#)
Tri-cities News. Diane Strandberg. Nov 2, 2018
- [Help Coquitlam, Port Moody teens save the bees](#)
Tri-cities News. Diane Strandberg. Feb 14, 2018