Jayden Teoh

Email ◇ GitHub ◇ LinkedIn ◇ Google Scholar

Professional Experience

Microsoft New York
Research Intern 05/2025 - Present
- Supervised by John Langford under the AI Frontiers group.
- Working on advancing the representation learning and planning capabilities of transformer models.

CARE.ai Labs Singapore
AI Researcher 05/2023 - Present
- Leading research projects focusing on deep Reinforcement Learning under guidance of Pradeep Varakantham.
- Published first author publications accepeted at NeurIPS 2024 (oral) and AAMAS 2024.

London Initiative for Safe AI (LISA) London
Research Assistant 06/2024 - 08/2024
- Research focusses on developing evaluations and elicitation mechanisms for AI sandbagging.
- Co-authored a publication accepted at ICML 2025.

Continental Singapore
Machine Learning Engineer Intern 04/2024 - 07/2024
- Developed a novel diffusion model for tire tread pattern generation.

Publications

On Generalization Across Environments In Multi-Objective Reinforcement Learning
Jayden Teoh, Pradeep Varakantham, Peter Vamplew View
- ICLR 2025

The Elicitation Game: Evaluating Capability Elicitation Techniques
Felix Hofstätter, Teun van der Weij, Jayden Teoh, Henning Bartsch, Francis Rhys Ward View
- ICML 2025

Improving Environment Novelty Quantification for Effective Unsupervised Environment Design
Jayden Teoh, Wenjun Li, Pradeep Varankantham View
- NeurIPS 2024 (oral)

Unifying Regret and State-Action Space Coverage for Effective Unsupervised Environment Design
Jayden Teoh, Wenjun Li, Pradeep Varankantham View
- AAMAS 2024 (Extended Abstract)

Education

Singapore Management University Singapore
Computer Science, Artificial Intelligence 2023-2026
- GPA >4.0/4.0. Ranked 1st in cohort.

Academic Service

- Reviewer for NeurIPS 2025 Datasets & Benchmarks Track
- Reviewer for ICLR 2025 Main Conference

Technical Skills

- **Programming Languages**: Python, Go, SQL, Javascript, Typescript, Java, C
- **ML/AI**: Mathematical Statistics, PyTorch, Reinforcement Learning, Natural Language Processing, Generative AI