

Jonathan Colaço Carr

School of Engineering and Applied Sciences
Harvard University
150 Western Ave
Allston, MA 02134

Email: jcolacocarr@g.harvard.edu

Education

Harvard University, Cambridge, MA

Ph.D. in Computer Science. Expected completion: June 2031.

Advisor: Milind Tambe

McGill University/Mila, Montreal, Canada

M.Sc. in Computer Science, August 2026.

Advisors: Prakash Panangaden and Doina Precup

Thesis title: *An analysis of optimization targets for value alignment within the Markov occupancy problem*

McGill University, Montreal, Canada

B.Sc. in Honours Mathematics, May 2023.

Minors in Computer Science and Physics

Honors and Awards

2026	FRQNT Doctoral Scholarship, \$100,000 Four-year scholarship for Canadian doctoral students, determined by academic record and research potential.
2024	MITACS Globalink Research Award, \$6,000 Research award for Canadian graduate students studying abroad, determined by expected research quality.
2024	McGill Graduate Mobility Award, \$8,900 Scholarship for McGill graduate students studying abroad, determined by academic standing.
2023	NSERC Canada Graduate Scholarship – Master’s, \$17,500 Graduate scholarship determined by academic excellence, research potential, and interpersonal skills.
2022	McGill Computer Science Undergraduate Research Award, \$7,000 Research award determined by academic record and research aptitude.
2021	McGill Space Institute Summer Undergraduate Research Award, \$7,000 Research award determined by academic record and extracurricular leadership.
2020	RBC x Microsoft AI for Social Impact Challenge, 2nd Place Prize, \$2,000 National AI competition prize determined by originality, technical excellence, impact, and feasibility.

Publications

* denotes co-first authorship.

Colaço Carr, J. (2026). “An analysis of optimization targets for value alignment within the Markov occupancy problem”. MA thesis. McGill University (in press).

Colaço Carr, J., P. Panangaden, D. Precup, and B. Van Roy (2026). *Reinforcement learning with pairwise preferences in long-term decision problems*. In: *International Conference on Machine Learning (ICML)*. acceptance rate: 26.6%.

Colaço Carr, J., Q. Sun, and C. Allen (2025). *Focused skill discovery: Learning to control specific state variables while minimizing side effects*. In: *Reinforcement Learning Journal (RLJ)*. acceptance rate: 39.0%.

Chehbouni^{*}, K., J. Colaço Carr^{*}, Y. More, J. C. Cheung, and G. Farnadi (2025). *Beyond the safety bundle: Auditing the Helpful and Harmless dataset*. In: *Proceedings of the Annual Conference of the Nations of the Americas Chapter of the ACL (NAACL)*. acceptance rate: 22.6%, **selected for an oral presentation**.

– Preliminary version selected for a spotlight presentation at the NeurIPS 2024 workshop on Algorithmic Fairness through the Lens of Metrics and Evaluations.

Colaço Carr, J., P. Panangaden, and D. Precup (2024). *Conditions on preference relations that guarantee the existence of optimal policies*. In: *International Conference on Artificial Intelligence and Statistics (AISTATS)*. acceptance rate: 27.6%.

Kennedy, J., J. Colaço Carr, S. Gagnon-Hartman, A. Liu, J. Mirocha, and Y. Cui (2024). *Machine-learning recovery of foreground wedge-removed 21-cm light cones for high-*z* galaxy mapping*. In: *Monthly Notices of the Royal Astronomical Society (MNRAS)*.

Research Experience

2026–Present **Harvard University**, Cambridge, MA

Student Researcher

Advisor: Milind Tambe

Researching efficient decision-making methods for HIV prevention.

2022–2026 **Mila – Quebec AI Institute**, Montreal, Canada

Student Researcher

Advisors: Prakash Panangaden and Doina Precup

Used human models from economics to design sequential decision-making algorithms aligned with human goals. Audited a human preference dataset for label quality, topic coverage, and changes in language-model behaviour after fine-tuning.

2024–2025 **Stanford University**, Stanford, CA

Visiting Student Researcher

Supervisor: Benjamin Van Roy

Extended the maximal lottery social choice rule to sequential decision-making problems and designed game-theory algorithms for computing solutions.

- 2024**
(Summer) **UC Berkeley (Center for Human-Compatible AI)**, Berkeley, CA
Supervisors: Cameron Allen and Stuart Russell
Research Intern
Developed a simple method to improve skill discovery using state abstraction, benefiting both efficiency and safety.
- 2021** **McGill Space Institute**, Montreal, Canada
Undergraduate Researcher
Supervisor: Adrian Liu
Implemented computer-vision algorithms to recover signals from noisy telescope data.

Work Experience

- 2024–2025** **Hortus AI**, Berkeley, CA
Lead Software Developer
Led a team of developers (15 in total) to build a consumer reports platform for AI tools. Reported directly to startup founder, Thomas Krendl Gilbert. Designed algorithms to rank and evaluate AI tools in collaboration with the GovAI coalition. Platform was used by over 100 government officials in its first three months of deployment.
- 2025**
(Fall) **The Democratic Engagement Exchange**, Toronto, Canada (remote)
Pilot Researcher
Conducted a feasibility study on the use of personalized AI assistants for voter education in Canada.
- 2024**
(Winter) **McGill University**, Montreal, Canada
Teaching Assistant for COMP 551: Applied Machine Learning
Held weekly office hours and created assignments for a graduate-level course in machine learning.
- 2022**
(Winter) **AltaML**, Toronto, Canada (remote)
Machine Learning Intern
Implemented machine learning algorithms to identify cross-selling opportunities.

Extracurricular Experience

- 2022–2023** **McGill AI Safety x Law Group**, Montreal, Canada
Co-founder
Co-founded and taught an eight-week course on AI safety for McGill Law students.
- 2019–2020** **RBC x Microsoft AI for Social Impact Challenge**, Montreal, Canada
Team Lead
Led a four-student team to design a textile recycling robot for an AI social impact challenge. Placed 2nd out of 167 teams from across Canada.
- 2019–2020** **The McGill Tribune**, Montreal, Canada
Editorial Board Member
Collaborated with journalists and graphic designers to publish articles for McGill's student newspaper.