

Carter Blair

Website | Google Scholar | carterblair@g.harvard.edu

EDUCATION

PhD in Computer Science

Harvard University

Cambridge, MA, USA

September 2025 – Present

- Supervised by Ariel Procaccia and Milind Tambe.

MMath in Computer Science

University of Waterloo

Waterloo, Canada

September 2023 – June 2025

- Supervised by Kate Larson and Edith Law.
- Thesis: [Deliberative Machines: From Reflective Dialogue to Fair Consensus with Language Models and Social Choice](#)

BSc in Computer Science and Psychology; Minor in Philosophy

University of Victoria

Victoria, Canada

Completed August 2022

- Supervised by Nishant Mehta and George Tzanetakis.
- Thesis: Online Positive Unlabelled Learning

PROFESSIONAL EXPERIENCE

Research Intern

[Vector Institute](#)

January 2024 – January 2025

Toronto, Canada

- Supervised by Gillian Hadfield.
- Developed AI-alignment methods that integrate existing normative infrastructure and institutions, contributing to two oral workshop papers at ICML 2024 and EC 2024.

Junior Data Scientist

[NannyML](#)

July 2022 – May 2023

Leuven, Belgium

- Developed scalable statistical methods for multivariate drift detection in production machine-learning systems and contributed them to NannyML's open-source monitoring library.

AWARDS AND HONORS

NSERC Doctoral Scholarship (PGS-D)

Natural Sciences and Engineering Research Council of Canada

2025 – 2028

\$120,000

- Selected for the CGS-D tier (highest-ranked applicants); accepted the internationally tenable PGS-D to hold the award at Harvard.

Cooperative AI PhD Fellowship (Scholar Tier)

Cooperative AI Foundation

2025 – 2028

- Selected as one of eight inaugural Scholars from 177 applicants.

President's Graduate Scholarship (Declined)

University of Waterloo

2025 – 2026

\$10,000

Research Dissemination Award

University of Waterloo

2024

\$500

Math Domestic Graduate Student Award

University of Waterloo

2023

\$2,000

Undergraduate Student Research Award

Natural Sciences and Engineering Research Council of Canada

2021

\$8,000

Dean's Entrance Scholarship

University of Victoria

2017

\$2,000

Conference Papers

5. **Blair, C.**, de Raaij, J., Procaccia, A. D., Rubin-Toles, M., Shah, N., Si, M., & Wang, S. (2026). The Structure of Bridging. In *Proceedings of the 27th ACM Conference on Economics and Computation (EC 2026)*, forthcoming. [Paper](#).
4. **Blair, C.**, & Larson, K. (2026). Generating Fair Consensus Statements with Social Choice on Token-Level MDPs. In *Proceedings of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026)*, pp. 1692–1700. [doi:10.65109/NZPR5925](#).
3. **Blair, C.**, Larson, K., & Law, E. (2025). Reflective Verbal Reward Design for Pluralistic Alignment. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence (IJCAI-25)* (Human-Centred AI track), pp. 10271–10279. [doi:10.24963/ijcai.2025/1141](#).
2. **Blair, C.**, Wang, X., & Perin, C. (2024). Quantifying Emotional Responses to Immutable Data Characteristics and Designer Choices in Data Visualizations. *IEEE VIS 2024*; published in *IEEE Transactions on Visualization and Computer Graphics*, 31(1), pp. 1006–1016. [doi:10.1109/TVCG.2024.3456361](#).
1. **Blair, C.**, Armstrong, B., & Larson, K. (2024). Liquid Ensemble Selection for Continual Learning. In *Proceedings of the 37th Canadian Conference on Artificial Intelligence (Canadian AI 2024)*. [doi:10.21428/594757db.d0ee747e](#).

Preprints

1. **Blair, C.**, Procaccia, A. D., & Tambe, M. (2026). Embeddings for Preferences, Not Semantics. [arXiv:2605.08360](#).

Peer-reviewed Workshop Papers, Posters, and Abstracts

5. Caiata, J., **Blair, C.**, & Larson, K. (2026). Procedural Fairness in Multi-Agent Bandits. Extended abstract in *Proceedings of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026)*, pp. 3531–3533. [doi:10.65109/RCLN5951](#).
4. **Blair, C.**, Armstrong, B., Alouf-Heffetz, S., Talmon, N., & Grossi, D. (2025). Probably Approximately Consensus: On the Learning Theory of Finding Common Ground. *2nd Workshop on Social Choice and Learning Algorithms (SCaLA) at IJCAI 2025*. [arXiv:2604.21811](#).
3. Trivedi, R., Chandak, N., **Blair, C.**, Sarkar, A., Weltman, T., Hadfield-Menell, D., & Hadfield, G. K. (2024). Altered Environments: The Role of Normative Infrastructure in AI Alignment. *Agentic Markets Workshop at ICML 2024*. Oral presentation.
2. Sarkar, A., Muresanu, A., **Blair, C.**, Sharma, A., Trivedi, R. S., & Hadfield, G. K. (2024). Normative Modules: A Generative Agent Architecture for Learning Norms that Supports Multi-Agent Cooperation. *Foundation Models and Game Theory Workshop at Economics and Computation (EC) 2024*. Oral presentation.
1. **Blair, C.**, & Wyeth, C. (2024). Decision Theoretic Planning with Language Models. *Foundation Models and Game Theory Workshop at Economics and Computation (EC) 2024*. Poster.

INVITED TALKS

- *Representative Preference Learning*, Google DeepMind, June 2024

TEACHING EXPERIENCE

Teaching Assistant, CS 480: Machine Learning <i>University of Waterloo</i>	September 2024 – December 2024
Teaching Assistant, CS 137: Programming Principles <i>University of Waterloo</i>	September 2023 – December 2023
Teaching Assistant, CSC 110: Fundamentals of Programming I <i>University of Victoria</i>	September 2021 – December 2021
Teaching Assistant, CSC 115: Fundamentals of Programming II <i>University of Victoria</i>	September 2020 – December 2020

PEER REVIEW

2027	AAAI Conference on Artificial Intelligence — Artificial Intelligence for Social Impact Track
2026	Conference on Neural Information Processing Systems (NeurIPS); IJCAI-ECAI Special Track on Human-Centred AI; 7th International Workshop on Agents for Societal Impact at AAMAS
2025	International Joint Conference on Artificial Intelligence (IJCAI); Social Choice and Learning Algorithms Workshop at IJCAI
2024	ACM Conference on Human Factors in Computing Systems (CHI) — Late-Breaking Work

LEADERSHIP AND SERVICE

Judge <i>National Collegiate Research Conference</i>	January 2026
Team Lead, University of Waterloo <i>Cooperative AI Foundation's Melting Pot Contest 2023 (NeurIPS 2023)</i>	September 2023 – November 2023
President <i>UVic Neuro-Tech Club</i>	September 2019 – April 2020
President <i>UVic Surf Club</i>	September 2018 – December 2019