

Sina Fazelpour

Assistant Professor,
Philosophy and Computer Science,
Northeastern University
Updated: September 17, 2026

s.fazel-pour@northeastern.edu
website: sinafazelpour.com
[Google Scholar Page](#)

Academic Positions

Assistant Professor, Northeastern University **2021-Present**
Department of Philosophy and Religion & The Khoury College of Computer Sciences

Past Positions

Visiting AI fellow, National Institute of Standards and Technology (NIST) **2022-2025**
Contributed to creating and operationalizing [NIST AI Risk Management Framework \(AI RMF\)](#)

Visiting Fellow, The Australian National University **January-March 2024**
Hosted by the School of Philosophy, College of Arts and Social Sciences

Visiting Scholar, UC Berkeley **June-August 2022**
Simons Institute for the Theory of Computing, Summer cluster “Interpretable Machine Learning”

Council Fellow, World Economic Forum **2020-2021**
The 2021 Council Fellow on the Global Future Council on Data Policy.

SSHRC Postdoctoral Fellow, Carnegie Mellon University **2019-2021**
Faculty Advisors: David Danks (CMU, Philosophy); Zachary C. Lipton (CMU, Machine Learning)

Graduate Research Fellow, W. Maurice Young Centre for Applied Ethics **2018-19**
Primary Investigator on the project “Trust and Conformity in Diverse Groups”.

Education

Ph.D., Philosophy. University of British Columbia **2019**
Faculty Advisors: Christopher Mole; Evan Thompson

B.A., Philosophy. University of Toronto **2013**
Awarded The Sanford Gold Medal for highest overall standing in the specialist program in philosophy.

M.Sc., Medical Biophysics. University of Toronto **2010**
Faculty Advisor: Christopher Macgowan

B.Eng., Electrical and Biomedical Engineering. McMaster University **2007**
With high distinction

Publications

Peer-reviewed articles

- [22] **Fazelpour, S.**, & Magnani, M. (Accepted) Aspirational Affordances of AI. *Philosophical Issues*, an annual journal supplement to *Noûs*.
- [21] **Fazelpour, S.**, O’Brien, J. & Rubin, H. (Accepted) Navigating Epistemic Monocultures in AI-Driven Science: A Simulation Study. *Philosophy of Science*
- [20] Gur-Arieh, S., Wang, A., & **Fazelpour, S.** (2026) Ambiguity Collapse by LLMs: A Taxonomy of Epistemic Risks. In *2026 ACM Conference on Fairness, Accountability, and Transparency (FAccT ’26)*.¹

¹Honorable Mention Best Paper Award.

- [19] Neumann, T., De-Arteaga, M., & **Fazelpour, S.** (2026, March). Should you use LLMs to simulate opinions? Quality checks for early-stage deliberation. In *Proceedings of the AAAI Conference on Artificial Intelligence*.²
- [18] **Fazelpour, S.** (2025). Disciplining Deliberation: A Socio-technical Perspective on Machine Learning Trade-Offs. *The British Journal for the Philosophy of Science*.
- [17] Akpınar, N. J. & **Fazelpour, S.**³ (2025). Authenticity and exclusion: social media recommendation algorithms and the dynamics of belonging in professional networks. *Synthese*.
- [16] Gao, Y., Lee, D., Burtch, G., & **Fazelpour, S.** (2025). Take Caution in Using LLMs as Human Surrogates. *Proceedings of the National Academy of Sciences*.
- [15] **Fazelpour, S.** & Fleisher, W. (2025). The value of disagreement in AI design, evaluation, and alignment. In *2025 ACM Conference on Fairness, Accountability, and Transparency (FAccT '25)*.
- [14] Neumann, T., De-Arteaga, M. & **Fazelpour, S.** (2022). Justice in misinformation detection systems: an analysis of algorithms, stakeholders, and potential harms. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*.⁴
- [13] **Fazelpour, S.**, & De-Arteaga, M. (2022). Diversity in sociotechnical machine learning systems. *Big Data & Society*, 9(1), 20539517221082027.
- [12] **Fazelpour, S.**, Lipton, Z. C., & Danks, D. (2022). Algorithmic fairness and the situated dynamics of justice. *Canadian Journal of Philosophy*, 52(1), 44-60.
- [11] **Fazelpour, S.**, & Rubin, H. (2022). Diversity and homophily in social networks. In *Proceedings of the 44th Annual Conference of the Cognitive Science Society*.
- [10] **Fazelpour, S.**, & Steel, D. (2022). Diversity, trust, and conformity: A simulation study. *Philosophy of Science*, 89(2), 209-231.
- [9] **Fazelpour, S.**, & Danks, D. (2021). Algorithmic bias: Senses, sources, solutions. *Philosophy Compass*, 16(8), e12760.
- [8] Dai, J., **Fazelpour, S.**, & Lipton, Z. (2021). Fair machine learning under partial compliance. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 55-65).⁵
- [7] **Fazelpour, S.** (2021). Norms in counterfactual selection. *Philosophy and Phenomenological Research*, 103(1), 114-139.
- [6] Steel, D., **Fazelpour, S.**, Crewe, B., & Gillette, K. (2021). Information elaboration and epistemic effects of diversity. *Synthese*, 198(2), 1287-1307.
- [5] Ransom, M., **Fazelpour, S.**, Markovic, J., Kryklywy, J., Thompson, E. T., & Todd, R. M. (2020). Affect-biased attention and predictive processing. *Cognition*, 203, 104370.
- [4] **Fazelpour, S.**, & Lipton, Z. C. (2020). Algorithmic fairness from a non-ideal perspective. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 57-63).
- [3] Steel, D., **Fazelpour, S.**, Gillette, K., Crewe, B., & Burgess, M. (2018). Multiple diversity concepts and their ethical-epistemic implications. *European journal for philosophy of science*, 8(3), 761-780.
- [2] Ransom, M., **Fazelpour, S.**, & Mole, C. (2017). Attention in the predictive mind. *Consciousness and Cognition*, 47, 99-112.
- [1] **Fazelpour, S.**, & Thompson, E. (2015). The Kantian brain: brain dynamics from a neurophenomenological perspective. *Current Opinion in Neurobiology*, 31, 223-229.

Book chapters

- [1] Ransom, M. & **Fazelpour, S.** (2021). Is attention just optimizing expected precision? in S. S. Gouveia, D. Mendonça, & M. Curado (Eds.) *The Philosophy and Science of Predictive Processing*. Bloomsbury Publishing.

²Runner-up, ISS Cluster Best Paper Award, INFORMS Information Systems Society, 2025.

³Equal contribution

⁴Article co-authored with a graduate student mentee.

⁵Article co-authored with an undergraduate student mentee.

Policy work and other

- [5] **Fazelpour, S.**, Hadfield-Menell, D., & Belli, L. (2024, February 9) Red teaming AI: The devil is in the details. *Tech Policy Press*. <https://www.techpolicy.press/red-teaming-ai-the-devil-is-in-the-details/>
- [4] Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST Trustworthy and Responsible AI, National Institute of Standards and Technology, Gaithersburg, MD, [online], <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>
- [3] Fogliato, R., **Fazelpour, S.**, Gupta, S., Lipton, Z., & Danks, D. (2022). Homophily and Incentive Effects in Use of Algorithms. arXiv preprint arXiv:2205.09701. Poster presentation at *the 44th Annual Conference of the Cognitive Science Society*.
- [2] World Economic Forum. (2021). [Pathways to Digital Justice \[White paper\]](#).
- [1] Jordan, S., **Fazelpour, S.**, Koshiyama, A., Kueper, J., DeChant, C., Leong, B., et al. (2019). Creating a tool to reproducibly estimate the ethical impact of artificial intelligence. *UCLA: The Program on Understanding Law, Science, and Evidence (PULSE)*. Retrieved from <https://escholarship.org/uc/item/56w756v8>

Under Review

- [2] **Fazelpour, S.**, & Rubin, H. Diversity and homophily in group inquiry.'
- [1] Wu, J., **Fazelpour, S.**, & Rubin, H. Modeling the Social Dynamics of Inquiry with Bayesian Agents.

Selected Talks

Invited talks

- [46] Governing AI in Science & Policy: Evaluating LLM-as-Surrogates. Keynote at University of Cincinnati's Center for Humanities and Technology Conference on Science and Technology in the Anthropocene (2026).
- [45] Aspirational Affordances of AI. Skilled Social Interactions in the Age of AI Workshop. University of Florida. (2025).
- [44] Sociotechnical Epistemology: Knowledge in an Algorithmic World. Keynote at Networks of (Mis)information Workshop. Utrecht University. (2025).
- [43] Sociotechnical Epistemology: Knowledge in an Algorithmic World. Computational Social Philosophy Seminars. (2025).
- [42] Aspirational Affordances of AI. Keynote at National Symposium on Equitable AI. Morgan State University. (2025).
- [41] Aspirational Affordances of AI. Microsoft Research NYC. (2025).
- [40] Authenticity & exclusion: social media algorithms and the dynamics of epistemic belonging. American Philosophical Association, Eastern Division. (2025).
- [39] AI trade-offs from a sociotechnical perspective. AI and the Human seminar series at Vanderbilt University. (2024)
- [38] Evaluating AI Capabilities as Affordances. University of Rochester. (2024).
- [37] Evaluating AI Capabilities as Affordances. Microsoft Research New England. (2024).
- [36] Bias in Generative AI. American Philosophical Association, Pacific Division. (2024). The Australian National University. (2024).
- [35] Evaluating AI Capabilities as Affordances. Workshop on Generative AI and Normative Philosophy of Computing. The Australian National University. (2024).
- [34] Trade-offs in AI systems. MINT Lab. The Australian National University. (2024).
- [33] Diversity of Human Factors in the Lifecycle of AI Systems. The School of Philosophy. The Australian National University. (2024).
- [32] Roundtable on Algorithmic Fairness. 2023 Public Interest Universities Network Convening. Boston University. (2023).

- [31] Using the NIST AI Risk Management Framework. IAPP Privacy. Security. Risk. San Diego. (2023).
- [30] Panel on data talent. The Impact Data Summit. The Mastercard Center for Inclusive Growth. (2023).
- [29] Trade-offs in sociotechnical ML systems. 2023 SSC Annual Meeting. Ottawa. (2023).
- [28] On interpreting formal trade-offs in sociotechnical ML systems. AI Metrology Colloquia Series. NIST. (2023).
- [27] Fairness in Sociotechnical Machine Learning Systems. Cornell Tech. (2023).
- [26] Applying the NIST AI Risk Management Framework to Inclusive Hiring (facilitator). The Partnership on Employment & Accessible Technology. (2023).
- [25] Justice in sociotechnical machine learning systems. William & Mary. (2023).
- [24] Diversity in AI. Penn Georgetown Digital Ethics Workshop. (2023).
- [23] Engineering as activism. Engineering Science Education Conference. University of Toronto. (2023).
- [22] Fairness in sociotechnical machine learning systems, Formal and Experimental Epistemology Conference (FAX2). University of Pittsburgh. (2022).
- [21] Accuracy and interpretability through the lens of human-AI teaming. Simons Institute for Theoretical Computing. UC Berkeley. (2022).
- [20] Where are the missing humans? Evaluating AI decision support systems in context. University of Toronto. Trust and the Ethics of AI (2022)
- [19] Disciplining deliberation: the case for exercising caution in interpreting formal trade-offs. Machine Wisdom Workshop 2. University of Pittsburgh. (2022).
- [18] Diversity in sociotechnical machine learning Systems. Boston University (2022). Care-AI, University of Guelph (2022). University of Memphis. (2022).
- [17] Evaluating algorithm-based decisions in context. Cognition, Values and Behaviour Lab Meetings series. Ludwig Maximilian University of Munich. (2021).
- [16] Socializing algorithms. DesignIt. (2021).
- [15] The epistemic benefits of group diversity. Northeastern Epistemology Workshop. Northeastern University. (2021).
- [14] Diversity, trust and conformity. Simulations of Scientific Inquiry (DFG Network), Munich Center for Mathematical Philosophy. (2021).
- [13] Diversity in sociotechnical machine learning systems. Machine Wisdom Workshop. University of Pittsburgh. (2021).
- [12] Algorithmic justice in a complex world. Symposium on Equity & Artificial Intelligence. University of Southern California. (2021).
- [11] The epistemic benefits of demographic diversity. University of Illinois at Chicago. (2021).
- [10] Algorithmic fairness in a complex world. Northeastern University. (2021).
- [9] Addressing algorithmic bias: Overcoming myopia and solutionism. Ethics of AI Conference. University of Texas at Austin. (2020).
- [8] Algorithmic bias: an overview. Guest lecture at University of British Columbia. (2020).
- [7] Algorithmic bias: senses, sources, and solutions. Guest lecture at University of Toronto. (2020).
- [6] Algorithmic fairness from a non-ideal perspective. Guest lecture at Australian National University. (2020).
- [5] Evaluating artificial intelligence decision support tools as embedded systems. Simons Institute for the Theory of Computing. University of California, Berkeley. (2020).
- [4] A non-ideal perspective on algorithmic fairness. Workshop on Algorithmic Ethics: Philosophy and Computer Science. The Goergen Institute for Data Science and the Department of Philosophy. University of Rochester. (2020).
- [3] Algorithmic fairness from a non-ideal perspective. Machine learning research group. CapitalOne. (2020).
- [2] The epistemic impact of diversity. The W. Maurice Young Centre for Applied Ethics. University of British Columbia. (2019).
- [1] Normality and counterfactuals: a functional approach. University of British Columbia. (2018).

Refereed Talks

- [23] Ambiguity Collapse by LLMs: A Taxonomy of Epistemic Risks. ACM Conference on Fairness, Accountability, and Transparency. (2026). Presented by co-author.
- [22] The Value of Disagreement in AI Design, Evaluation, and Alignment. ACM Conference on Fairness, Accountability, and Transparency. (2025). Presented by co-author.
- [21] Aspirational Affordances of AI. Penn-Georgetown Digital Ethics Workshop. (2025).
- [20] Authenticity & exclusion: social media algorithms and the dynamics of epistemic belonging. MINT-Yale Law School Workshop on Normative Philosophy of Computing. (2024).
- [19] Practice tutorial: Using the NIST AI Risk Management Framework. ACM Conference on Fairness, Accountability, and Transparency. (2023).
- [18] Resisting algorithmic homogenization: the value of disagreement in data labeling. Philosophy, AI and Society (PAIS) Network Workshop, Stanford University. (2023).
- [17] Justice in Misinformation Detection Systems: An Analysis of Algorithms, Stakeholders, and Potential Harms. ACM Conference on Fairness, Accountability, and Transparency. (2022). Presented by co-author.
- [16] Translation tutorial: Sociocultural diversity in machine learning: Lessons from philosophy, psychology, and organizational science. ACM Conference on Fairness, Accountability, and Transparency. (2021).
- [15] Fairness under partial compliance: A simulation study. ACM Conference on AI, Ethics, and Society (AIES'21). Presented by co-author. (2021).
- [14] Algorithmic fairness from a non-ideal perspective. American Philosophical Association, Eastern Division. (2021).
- [13] Diversity, trust, and conformity. East European Network for Philosophy of Science. (2020). Postponed due to the pandemic.
- [12] Algorithmic injustice from a non-ideal perspective. Canadian Society for the History and Philosophy of Science. (2020). Postponed due to the pandemic.
- [11] A non-ideal perspective on algorithmic fairness. Ethics of Data Science Conference. (2020). Cancelled due to the pandemic.
- [10] Algorithmic fairness from a non-ideal perspective. Philosophy of Science Association. (2020). Postponed due to the pandemic.
- [9] Algorithmic fairness from a non-ideal perspective. ACM Conference on AI, Ethics, and Society (AIES'20). (2020).
- [8] The aspirations and shortcomings of causal approaches to algorithmic fairness. 10th Workshop in Decisions, Games, and Logic: Ethics, Statistics, and Fair AI. Caltech. (2019).
- [7] Blind spots and models: The case of information elaboration in philosophical models of diversity. Canadian Society for the History and Philosophy of Science. (2018).
- [6] The Janus face of counterfactual thinking. Interdisciplinary Workshop on Counterfactual Thinking. University of Toronto. (2016).
- [5] Three problems for the predictive coding theory of attention. Minds Online Conference. (2015).
- [4] The predictive coding model of dreaming. 15th Congress of Logic, Methodology and Philosophy of Science. Helsinki, Finland. (2015).
- [3] Bayesian dreams. International Society for the History, Philosophy, and Social Studies of Biology. Montreal, Canada. (2015).
- [2] Three problems for the predictive coding theory of attention. Society for Philosophy and Psychology annual meeting. Presented by co-author. Honorable mention for the William James Prize. (2015).
- [1] The saliency of pain: A structural account of pain affect. The 20th Toward a Science of Consciousness. Tucson, Arizona. (2014).

Grants

External Grants

- Survival and Flourishing Fund, Speculation Grant. “Computational Social Philosophy” (PI). \$40,000 USD. 2026.
- Alfred P. Sloan Foundation. “Cultivating Robust AI and Society Research Capacity and Community” (Co-Investigator, PI Kathleen Creel). \$400,000 USD. December 2025–December 2026.
- Survival and Flourishing Fund, Speculation Grant,. “Socially Situated Evaluation of AI Risks” (PI). \$30,000 USD. 2025.
- Schmidt Sciences AI2050 ECF Collaboration Fund. “Exploring AI’s Influence on Perceptions of Future Possibilities” (co-PI). \$29,975 USD. 2025.
- Schmidt Sciences AI2050 ECF Collaboration Fund. “Philosophy of AI Research Convening” (co-PI). \$27,816 USD. 2025.
- Schmidt Sciences AI2050 ECF Collaboration Fund. “The Philosophy of Artificial Intelligence Network Talks”. (co-PI). \$43,241 USD. 2024.
- Schmidt Futures’ AI2050 ECF Collaboration Fund. “Operationalizing NIST’s AI Risk Management Framework: Mapping the Human Factor”. \$30,000 USD. 2023.
- Schmidt Sciences AI 2050 Early Career Fellowship. “An integrated prescriptive framework for human-AI collaboration” (PI). \$242,517 USD. 2023.
- Omidyar Network. Restricted gifts to the IDEAS summer program. \$200,000. 2023.
- New America’s Public Interest Technology University Network (“PIT-UN”) Challenge. “Intelligence, Data, Ethics and Society” (PI). \$89,645 USD. 2022.
- National Science Foundation. “Research Community Development Grant: Summer Training Program to Expand the AI and Data Ethics Research Community” (Senior Key personnel; PI Ron Sandler). \$299,374 USD. 2022-2023.
- Amazon Research Awards. “An Integrated Framework for Understanding Human-AI Hybrid Decision-Making” (co-applicant; PI David Danks). 2021. \$63,000 USD in funding and additional \$40,000 USD in AWS Promotional Credits.
- Social Sciences and Humanities Research Council Connection Grant. “Causation and Counterfactuals in Philosophy, Science and the Law” (written with PI Paul Bartha). 2018, \$16,354 CAD
- Templeton Foundation, Summer Seminars in Neuroscience and Philosophy. “Attention in the Predictive Mind” (co-PI Rebecca Todd, Madeleine Ransom, Jelena Markovic). 2016, \$25,000 USD

Internal Grants

- NuLab Seedling Grant, Northeastern University. “Fair explorations: The structure and dynamics of responsible research & innovation communities” (PI). 2021. \$2,500 USD.
- Block Center for Technology & Society, Carnegie Mellon University. “An Integrated Framework for Studying and Regulating Human-AI Hybrid Decision-Making Systems” (co-applicant; PI Zachary Lipton). 2020, \$60,000 USD

Fellowships and Awards

Fellowships and Scholarships

- Re-imagining/Re-forming Fellowship, Humanities Center, Northeastern University, 2022-2023
- Social Sciences and Humanities Research Council of Canada’s Postdoctoral Fellowship, 2019-2021, \$90,000 CAD
- W. Maurice Young Centre for Applied Ethics Graduate Fellowship, 2018-2019, \$20,000 CAD
- Public Scholars Initiative Research Fellowship, University of British Columbia, 2018, \$4,500 CAD

- Public Scholars Initiative Fellowship, University of British Columbia, 2017, \$8,500 CAD
- Joseph-Armand Bombardier Canada's Doctoral Scholarship, 2014-2017, \$105,000 CAD
- University of British Columbia Four-Year Fellowship, 2013-2017, \$72,000 CAD
- Victoria College, The Regents In-Course Scholarship, 2012, \$1,000 CAD
- Research Training Competition Fellowship, The Hospital for Sick Children, \$40,000 CAD
- Ontario Graduate Scholarship, 2008-9 (Declined), \$30,000 CAD

Awards and Honors

- Best Paper Honorable Mention, ACM Conference on Fairness, Accountability, and Transparency, 2026 (Gur-Arieh, Wang & Fazelpour).
- Runner-up, ISS Cluster Best Paper Award, INFORMS Information Systems Society, 2025 (Neumann, De-Arteaga & Fazelpour, "Should you use LLMs to simulate opinions?").
- Top Cited Article 2021–2022, *Philosophy Compass* (Wiley), for "Algorithmic bias: Senses, sources, solutions."
- Third PSA Women's Caucus Prize Symposium for organizing the "Conceptual and Methodological Challenges in Algorithmic Fairness" symposium, 2020 (conference postponed to 2021 due to COVID)
- International Policy Ideas Challenge, Global Affairs Canada, 2020
- Don Brown Graduate Teaching Award, 2018
- UBC Ambassador for Congress of the Humanities and Social Sciences, 2018, \$1,500 CAD
- Honorable mention for the William James Prize, Society for Philosophy and Psychology, 2016
- The Sanford Gold Medal in Specialist Program in Philosophy, Victoria College, Toronto, 2013
- Dean's list, University of Toronto, 2011-2013
- The Walter and Mary Tuohy Award, 2012, \$2,000 CAD
- PEM Research in Progress Seminar Award, The Hospital for Sick Children
- Dean's list, McMaster University, 2004-2007
- The Dr. Harry Lyman Hooker Scholarship Award for Academic Excellence, 2005, 2006, \$1,500 CAD
- Natural Sciences and Engineering Research Council Award for Undergraduate Summer Research, 2005, \$6,000 CAD
- Member of The National Organization for Development of Exceptional Talents (NODET), Iran, 1994-2001

Teaching and Advising

Teaching Awards and Recognitions

Don Brown Graduate Teaching Award, 2018

Courses

Northeastern University

Graduate courses

S23-24 CS8982. Directed reading on the ethics of human-AI collaboration and teaming.

F23-24 CS48982. Directed reading on the ethics of human-AI collaboration and teaming.

Undergraduate courses

F25-26 AI & Society. PHIL/CS4050 (15 students). Cross-listed between Computer Science & Philosophy.

S24-25 AI & Society. PHIL/CS4050 (19 students). Cross-listed between Computer Science & Philosophy.

F23-24 Values & Sociotechnical Algorithmic Systems. PHIL/CS4050 (18 students). Cross-listed between Computer Science & Philosophy.

F22-23 Technology & Human Values. PHIL 1145 (40 students). Department of Philosophy & Religion.

F22-23 Philosophy of Mind. PHIL 4535 (19 students). Department of Philosophy & Religion.
S21-22 Technology & Human Values. PHIL 1145: Sections 3 (40 students) and 6 (40 students). Department of Philosophy & Religion.

Carnegie Mellon University

Graduate seminar

S20-21 Ethics & Policy of Data Analytics. 94-836 (20 students). Heinz College.

University of British Columbia

Undergraduate courses

S18-19 Understanding and Designing Cognitive Systems. COGS300 (60 students). Cognitive Systems.

S17-18 Understanding and Designing Cognitive Systems. COGS300 (72 students). Cognitive Systems.

F16-17 Introduction to Scientific Reasoning. PHIL125 (24 students). Philosophy Department.

Short courses

S17-18 Causal Graphical Models: Representation and Inference. University of British Columbia.

Teaching Assistant, University of British Columbia

Teaching assistant for various courses at different levels, including Introduction to Philosophy (PHIL101) (×2), Introduction to Ethics (PHIL102), Symbolic Logic (PHIL220) (×3), Philosophy of Law (PHIL338) (×2), Biomedical Ethics (PHIL433), Philosophy of Mind (PHIL451). Advanced Topics in Philosophy of Mind (PHIL491). Winter 2014-Fall 2017.

Teaching training

Certificate Program in Advanced Teaching and Learning at University of British Columbia.

Instructional Skills Workshop, UBC Graduate Pathways to Success.

Service and Professional Development

Service to the Profession

- Steering Committee, for [ACM Conference on Fairness, Accountability, and Transparency](#). Term: 2024-2028.
- Organizer, Intelligence, Data, Ethics and Society (IDEAS) summer workshop, Boston. 2021-present. See: [the IDEAS Boston 2021 website](#).
- Organizer, [Computational Social Philosophy Seminars \(CSPS\)](#), a bi-weekly online talk series. 2024-present.
- Position Paper Area Chair for [NeurIPS 2026](#).
- Program Chair for [ACM Conference on Fairness, Accountability, and Transparency 2025](#).
- Organizer, Philosophy of AI Network Talks (PAINT), a bi-weekly online talk series. 2025-present. See: [PAINT series website](#).
- Philosophy Area Chair [ACM Conference on Fairness, Accountability, and Transparency 2024](#).
- Tutorial Co-chair [ACM Conference on Fairness, Accountability, and Transparency 2023](#).
- Review panelist, [The Dangers and Opportunities of Technology: Perspectives from the Humanities program](#), National Endowment for the Humanities Office of Digital Humanities. 2023.
- Reviewer, *Science*, *Science and Engineering Ethics*, *Politics, Philosophy & Economic*, *Nature Human Behaviour*, *Philosopher's Imprint*, *Philosophy of Science*, *The British Journal for the Philosophy of Science*, *Synthese*, *Studies in History and Philosophy of Science*, *AI & Society*, *AAAI/ACM Conference on Artificial Intelligence*,

Ethics, and Society, NeurIPS, ACM Conference on Fairness, Accountability, and Transparency, Minds online, Dialogue: Canadian Philosophical Review, Frontiers in Psychology.

- Symposium organizer, “Conceptual and Methodological Challenges in Algorithmic Fairness” at *The 27th Biennial Meeting of the Philosophy of Science Association*. Awarded the [Third PSA Women’s Caucus Prize Symposium](#).
- Program Committee, NeurIPS Workshop on Fair AI in Finance

Service to the Community

- Speaker, Science on Screen introduction to *Minority Report* (Talk on AI in criminal justice), Coolidge Corner Theatre, Boston. 2022.
- Speaker, Measuring fairness in AI. Tapestry Retirement Home At Wesbrook Village, Vancouver. 2018.

Languages

Persian (native speaker), English (fluent)