

Sina Fazelpour

Assistant Professor,
Philosophy and Computer Science,
Northeastern University

s.fazel-pour@northeastern.edu
website: sinafazelpour.com

Updated: August 5, 2022

Academic Positions

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

Past Positions

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

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

Graduate Research Fellow, W. Maurice Young Centre for Applied Ethics **2018-19**
Primary Investigator on the project “Trust and Conformity in Diverse Groups: a Simulation-based Investigation”, carried out at the University of British Columbia’s School of Population and Public Health.

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

- [16] **Fazelpour, S.** & Rubin, H. (2022). Diversity and homophily in social networks. In J. Culbertson, A. Perfors, H. Rabagliati, & V. Ramenzoni (Eds.), *Proceedings of the 44th annual meeting of the cognitive science society*.
- [15] Fogliato, R., **Fazelpour, S.**, Gupta, S., Lipton, Z., & Danks, D. (2022). Homophily and incentive effects in use of algorithms. In J. Culbertson, A. Perfors, H. Rabagliati, & V. Ramenzoni (Eds.), *Proceedings of the 44th annual meeting of the cognitive science society*.
- [14] Neumann, T., De-Arteaga, M. & **Fazelpour, S.** (2022). Justice in misinformation detection systems: an analysis of algorithms, stakeholders, and potential harms. In *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAccT)*.
- [13] **Fazelpour, S.** & De-Arteaga, M. (2022) Diversity in sociotechnical machine learning systems. *Big Data & Society*.

- [12] **Fazelpour, S.** & Steel, D. (2022) Diversity, trust, conformity: a simulation study. *Philosophy of Science*.
- [11] **Fazelpour, S.** & Danks, D. (2021). Algorithmic bias: Sources, senses, solutions. *Philosophy Compass*.
- [10] **Fazelpour, S.**, Lipton, Z., & Danks, D. (2021). Algorithmic fairness and the dynamics of justice. *Canadian Journal of Philosophy* in special issue on “The Political Philosophy of Data and AI”.
- [9] Dai, J., **Fazelpour, S.**, Lipton, Z. (2021). Fairness under partial compliance: A simulation study. In *2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES’21)*.
- [8] **Fazelpour, S.** (2020). Norms in counterfactual selection. *Philosophy and Phenomenological Research*.
- [7] Ransom, M., **Fazelpour, S.**, Markovic, J., Kryklywy, J., Thompson, E. T., & Todd, R. M. (2020). Affect-biased attention and predictive processing. *Cognition*.
- [6] **Fazelpour, S.** & Lipton, Z. C. (2020). Algorithmic fairness from a non-ideal perspective. In *2020 AAAI/ACM Conference on AI, Ethics, and Society (AIES’20)*.
- [5] Steel, D. **Fazelpour, S.**, Crewe, B. & Gillette, K. (2019). Information elaboration and epistemic effects of diversity. *Synthese*.
- [4] 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*.
- [3] Ransom, M., **Fazelpour, S.**, & Mole, C. (2017). Attention in the predictive mind. *Consciousness and Cognition*.
- [2] **Fazelpour, S.**, & Thompson, E. (2015). The Kantian brain: brain dynamics from a neurophenomenological perspective. *Current Opinion in Neurobiology*.
- [1] Hojjat, S. P., **Fazelpour, S.**, & Shirani, S. (2007). Multiple description coding of video using phase scrambling. In *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*.

Invited Publications

- 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.
- 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>

Policy white papers

- World Economic Forum. (2021). Pathways to Digital Justice [White paper]. <https://www.weforum.org/whitepapers/pathways-to-digital-justice>.

Fellowships and Awards

Fellowships and Scholarships

- Re-imagining/Re-forming Fellowship, Humanities Center, Northeastern University, 2022-2023
- Social Sciences and Humanities Research Council 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 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

Grants

NuLab Seedling Grant. “Fair explorations: The structure and dynamics of responsible research & innovation communities” (PI). 2021. \$2,500 USD.

Amazon Research Awards. “An Integrated Framework for Understanding Human-AI Hybrid Decision-Making” (co-applicant with David Danks, Zachary Lipton). 2021. \$63,000 USD in funding and additional \$40,000 USD in AWS Promotional Credits.

Block Center for Technology & Society, Carnegie Mellon University. “An Integrated Framework for Studying and Regulating Human-AI Hybrid Decision-Making Systems” (co-applicant with Zack Lipton, David Danks). 2020, \$60,000 USD

Social Sciences and Humanities Research Council Connection Grant. “Causation and Counterfactuals in Philosophy, Science and the Law” (written with 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

Awards and Honours

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

NSERC Award for Undergraduate Summer Research, 2005, \$6,000 CAD

Member of The National Organization for Development of Exceptional Talents (NODET), Iran, 1994-2001

Selected Talks

Invited talks

Accuracy and interpretability through the lens of human-AI teaming. Simons Institute for Theoretical Computing. UC Berkeley. (2022).

Where are the missing humans? Evaluating AI decision support systems in context. University of Toronto. Trust and the Ethics of AI (2022)

Disciplining deliberation: the case for exercising caution in interpreting formal trade-offs. Machine Wisdom Workshop 2. University of Pittsburgh. (2022).

Diversity in sociotechnical machine learning Systems. Boston University (2022). Care-AI, University of Guelph (2022). University of Memphis. (2022).

Evaluating algorithm-based decisions in context. Cognition, Values and Behaviour Lab Meetings series. Ludwig Maximilian University of Munich. (2021).

Socializing algorithms. DesignIt. (2021).

The epistemic benefits of group diversity. Northeastern Epistemology Workshop. Northeastern University. (2021).

Diversity, trust and conformity. Simulations of Scientific Inquiry (DFG Network), Munich Center for Mathematical Philosophy. (2021).

Diversity in sociotechnical machine learning systems. Machine Wisdom Workshop. University of Pittsburgh. (2021).

Algorithmic justice in a complex world. Symposium on Equity & Artificial Intelligence. University of Southern California. (2021).

The epistemic benefits of demographic diversity. University of Illinois at Chicago. (2021).

Algorithmic fairness in a complex world. Northeastern University. (2021).

Addressing algorithmic bias: Overcoming myopia and solutionism. Ethics of AI Conference. University of Texas at Austin. (2020).

Algorithmic bias: an overview. Guest lecture at University of British Columbia. (2020).

Algorithmic bias: senses, sources, and solutions. Guest lecture at University of Toronto. (2020).

Algorithmic fairness from a non-ideal perspective. Guest lecture at Australian National University. (2020).

Evaluating artificial intelligence decision support tools as embedded systems. Simons Institute for the Theory of Computing. University of California, Berkeley. (2020).

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).

Algorithmic fairness from a non-ideal perspective. Machine learning research group. CapitalOne. (2020).

The epistemic impact of diversity. The W. Maurice Young Centre for Applied Ethics. University of British Columbia. (2019).

Normality and counterfactuals: a functional approach. University of British Columbia. (2018).

Referred Talks

Fairness under partial compliance: A simulation study. ACM Conference on AI, Ethics, and Society (AIES'21). Presented by co-author. (2021).

Algorithmic fairness from a non-ideal perspective. American Philosophical Association, Eastern Division. (2021).

Diversity, trust, and conformity. East European Network for Philosophy of Science. (2020). Postponed due to the pandemic.

Algorithmic injustice from a non-ideal perspective. Canadian Society for the History and Philosophy of Science. (2020). Postponed due to the pandemic.

A non-ideal perspective on algorithmic fairness. Ethics of Data Science Conference. (2020). Cancelled due to the pandemic.

Algorithmic fairness from a non-ideal perspective. Philosophy of Science Association. (2020). Postponed due to the pandemic.

Algorithmic fairness from a non-ideal perspective. ACM Conference on AI, Ethics, and Society (AIES'20). (2020).

The aspirations and shortcomings of causal approaches to algorithmic fairness. 10th Workshop in Decisions, Games, and Logic: Ethics, Statistics, and Fair AI. Caltech. (2019).

Blind spots and models: The case of information elaboration in philosophical models of diversity. Canadian Society for the History and Philosophy of Science. (2018).

The Janus face of counterfactual thinking. Interdisciplinary Workshop on Counterfactual Thinking. University of Toronto. (2016).

Three problems for the predictive coding theory of attention. Minds Online Conference. (2015).

The predictive coding model of dreaming. 15th Congress of Logic, Methodology and Philosophy of Science. Helsinki, Finland. (2015).

Bayesian dreams. International Society for the History, Philosophy, and Social Studies of Biology. Montreal, Canada. (2015).

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).

The saliency of pain: A structural account of pain affect. The 20th Toward a Science of Consciousness. Tucson, Arizona. (2014).

Teaching Experience

Teaching Awards and Recognitions

Don Brown Graduate Teaching Award, 2018

Instructor

Northeastern University

Undergraduate courses

F22-23 Technology & Human Values (PHIL 1145).

F22-23 Philosophy of Mind (PHIL 4535).

S21-22 Technology & Human Values (PHIL 1145). Sections 3 and 6.

Carnegie Mellon University

Graduate seminar

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

University of British Columbia

Undergraduate courses

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

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

F16-17 Introduction to Scientific Reasoning (PHIL125). 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.

Professional Services

Organizer, Intelligence, Data, Ethics and Society (IDEAS) summer workshop (2021-)

Reviewer, *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*.

Program Committee, NeurIPS Workshop on Fair AI in Finance
Symposium organizer, “Conceptual and Methodological Challenges in Algorithmic Fairness” at *The 27th Biennial Meeting of the Philosophy of Science Association*
Organizer, Ethics & Artificial Intelligence Reading Group, Carnegie Mellon University, 2019-Present
UBC Ambassador for Congress of the Humanities and Social Sciences Meetings 2018 and 2019
Organizer, Trust and Transparency in Machine Learning Systems Reading Group, University of British Columbia, 2018-2019
Co-organizer, Workshop on Causation and Counterfactuals in Philosophy, Science, and Law, University of British Columbia, 2018
Organizer, Graduate Student Colloquium Series, Philosophy Department, University of British Columbia, 2016

Languages

Persian (native speaker), English (fluent)

Skills

Matlab, Python, Agent-based modeling, Time-frequency analysis, and R (basics).