Amir H. Karimi

ASSISTANT PROFESSOR OF MACHING LEARNING | ML SCIENTIST & ENGINEER

✉ amirhkarimi@gmail.com | 🏠 www.amirhkarimi.com | 🐦 amirhkarimi_

Employment

University of Waterloo & Vector Institute
ASSISTANT PROFESSOR OF MACHINE LEARNING
Waterloo & Toronto, CA
Sep 2023 - p.
- Tenure-track Assistant Professor in the Electrical and Computer Engineering Department

Google DeepMind
RESEARCH SCIENTIST INTERN
London, UK
May 2022 - Oct 2022
- Improved search via type- & IO-based neurally-guided program synthesis. Mentors: Lars Buesing, David Amos, Jessica Hamrick

Google Brain
RESEARCH SCIENTIST INTERN
Waterloo, CA
Dec 2021 - Apr 2022
- Investigated the causal effect of training hyperparameters on ML explanation. Mentors: Been Kim, Simon Kornblith
- Successfully published at ICML 2023, under the title *"On the Relationship Between Explanation and Prediction: A Causal View"* [C16]

Meta (Facebook) Inc.
SOFTWARE ENGINEER
New York, USA
Aug 2015 - Sep 2016
- Full-stack software engineer on the Enterprise Eng. team, responsible for front-end dev using React and hphp among others
- Implemented the landing page, a customizable notification framework, and the testing and logging platform for the Org Tool
- Successfully published at EMNLP 2016, under the title *"Key-Value Memory Networks for Directly Reading Documents"* [C2, P1]

Meta (Facebook) Inc.
SOFTWARE ENGINEER INTERN
Menlo Park, USA
Feb 2014 - Apr 2014
- Delivered a first-of-a-kind reporting tool for Facebook's largest Business Manager ad clients to view and manage historical ad budgets
- Successfully reduced TTI < 7sec for Facebook's largest ad clients (100K+ ad accounts), by optimizing front-end JavaScript rendering

BlackBerry Inc.
SOFTWARE ENGINEER INTERN
Toronto, CA
May 2013 - Dec 2013
- Implemented an automation framework to test 6,000+ ported selenium webdriver and WebKit tests for BB10
- Successfully integrated browser team's test automation results with company central test database

Stanford University
UNDERGRADUATE RESEARCH ASSISTANT - HIGH-FREQUENCY LAB
Stanford, USA
May 2012 - Aug. 2012
- Developed a helical antenna and used machine learning for precise 2D localization enabling hand gesture recognition.

Education

Max Planck Institute for Intelligent Systems & ETH Zurich
PHD IN COMPUTER SCIENCE
Tübingen, DE & Zürich, CH
Oct 2018 - Jul 2023
- **Thesis**: "Advances in Algorithmic Recourse: Ensuring Causal Consistency, Fairness, & Robustness"
- **Supervisors**: Prof. Bernhard Schölkopf & Prof. Isabel Valera
- **Major Awards**: NSERC CGS-D PhD Fellowship, Max Planck ETH PhD Fellowship, Google PhD Fellowship, ETH Zurich Medal

University of Waterloo
MMATH IN COMPUTER SCIENCE
Waterloo, CA
Sep 2016 - Apr 2018
- **Thesis**: "Exploring New Forms of Random Projections for Dimensionality Reduction"
- **Supervisors**: Prof. Alexander Wong & Prof. Ali Ghodsi
- **Major Award**: *Alumni Gold Medal* for highest standing across all master's programs at UWaterloo

University of Toronto
B.A.SC. IN ENGINEERING SCIENCE – ELECTRICAL AND COMPUTER STREAM
Toronto, CA
Sep 2010 - Jun 2015
- **Thesis**: "Benchmarking a Neuro-biologically Inspired Adaptive Controller"
- **Supervisors**: Prof. Chris Eliasmith & Prof. Richard Zemel
- **Major Award**: *Spirit of Engineering Science Award* for outstanding community contribution

Publication Highlights

SPOTLIGHT (💡)   ORAL (🎤)   ≥ 100 CITATIONS (☆)   BEST PAPER (🏆)   PATENT (✳)   BOOK CHAPTER (📖)   EQUAL CONTRIBUTION (*)

|💡|	|🎤|	|☆|	|🏆|	|✳|	|📖|		H-INDEX	G-INDEX	MAX CITATIONS	TOTAL CITATIONS
3	10	5	1	1	1		14	29	1,173 [C2]	2,816

Under Preparation

UP2	-	2024	-	-	*"Imagining and building wise machines: The centrality of AI metacognition"* Johnson, **Karimi**, Bengio, Chater, Gerstenberg, Larson, Levine, Mitchell, Rahwan, Schölkopf, Grossmann
UP1	-	2024	-	-	*"Explainable AI is Causal Discovery in Disguise"* **Karimi**

Patents

P1	✸	2018	*US Patent*	-	*"Key-Value Memory Networks"* Miller, Fisch, Dodge, **Karimi**, Bordes, Weston

Book Chapters

B1	📙	2022	*Springer LNAI*	-	*"Towards Causal Algorithmic Recourse"* **Karimi***, von Kügelgen*, Schölkopf, Valera

Journal Proceedings

J2	☆	2022	*ACM Comp. Surveys*	I.F. 23.8	*"A survey of algorithmic recourse: contrastive explanations & consequential …"* **Karimi**, Barthe, Schölkopf, Valera
J1	-	2011	*Optics Express*	I.F. 3.2	*"Automated detection and density assessment of keratocytes in the human …"* **Karimi**, Wong, Bizheva

Conference Proceedings

C18	-	2024	*ICML*	A.R. %27.5	*"Prospector Heads: Generalized Feature Attribution for Large Models & Data"* Machiraju, Derry, Desai, Guha, **Karimi**, Zou, Altman, Ré, Mallick
C17	🎤	2023	*AAAI*	A.R. %23.75	*"Causal Adversarial Perturbations for Individual Fairness and Robustness in …"* Ehyaei, Mohammadi, **Karimi**, Samadi, Farnadi
C16	-	2023	*ICML*	A.R. %27.9	*"On the Relationship Between Explanation and Prediction: A Causal View"* **Karimi**, Muandet, Kornblith, Schölkopf, Kim
C15	-	2023	*ICML*	A.R. %27.9	*"On Data Manifolds Entailed by Structural Causal Models"* Dominguez-Olmedo, **Karimi**, Arvanitidis, Schölkopf
C14	-	2023	*FAccT*	A.R. %24.6	*"Robustness Implies Fairness in Causal Algorithmic Recourse"* Ehyaei, **Karimi**, Schölkopf, Maghsudi
C13	💡	2022	*ICML*	A.R. %21.9	*"On the Robustness of Causal Algorithmic Recourse"* Dominguez-Olmedo, **Karimi**, Schölkopf
C12	🎤	2022	*AAAI*	A.R. %15.0	*"On the Fairness of Causal Algorithmic Recourse"* von Kügelgen, **Karimi**, Bhatt, Valera, Weller, Schölkopf
C11	🎤	2021	*ACM-AIES*	A.R. %38.0	*"Scaling Guarantees for Nearest Counterfactual Explanations"* Mohammadi, **Karimi**, Barthe, Valera
C10	☆ 💡	2021	*ACM-FAccT*	A.R. %25.0	*"Algorithmic Recourse: from Counterfactual Explanations to Interventions"* **Karimi**, Schölkopf, Valera
C9	☆ 💡	2020	*NeurIPS*	A.R. %20.1	*"Algorithmic recourse under imperfect causal knowledge: a probabilistic …"* **Karimi***, von Kügelgen*, Schölkopf, Valera
C8	☆ 🎤	2019	*AISTATS*	A.R. %32.4	*"Model-Agnostic Counterfactual Explanations for Consequential Decisions"* **Karimi**, Barthe, Balle, Valera
C7	🎤	2018	*IJCNN*	A.R. %22.7	*"Distance Correlation Autoencoder"* Wang, **Karimi**, Ghodsi
C6	🎤	2018	*CVIS*	A.R. ≥%40.0	*"FEELS: a full-spectrum enhanced emotion learning system for assisting …"* **Karimi***, Boroomand*, Pfisterer, Wong
C5	🏆 🎤	2017	*CVIS*	A.R. ≥%40.0	*"Ensembles of Random Projections for Nonlinear Dimensionality Reduction"* **Karimi**, Shafiee, Ghodsi, Wong
C4	-	2017	*CCN*	A.R. ≥%40.0	*"Synthesizing Deep Neural Network Architectures using Biological Synaptic …"* **Karimi**, Shafiee, Ghodsi, Wong
C3	🎤	2017	*ICIAR*	A.R. ≥%40.0	*"Discovery Radiomics via a Mixture Sequencers for Multi-Parametric MRI …"* **Karimi**, Chung, Shafiee, Khalvati, Haider, Ghodsi, Wong
C2	☆ 🎤	2016	*EMNLP*	A.R. %24.3	*"Key-Value Memory Networks for Directly Reading Documents"* Miller, Fisch, Dodge, **Karimi**, Bordes, Weston

C1	🎤	2016	*ICIP*	A.R. ≥%40.0	*"Spatio-temporal saliency detection using abstracted fully-connected …* **Karimi**, Shafiee, Scharfenberger, BenDaya, Haider, Talukdar, Clausi, Wong

Workshop Proceedings

W2	-	2018	*NeurIPS*	-	*"Deep Variational Sufficient Dimensionality Reduction"* Banijamali, **Karimi**, Ghodsi
W1	-	2017	*NeurIPS*	-	*"JADE: Joint Autoencoders for Dis-Entanglement"* **Karimi***, Banijamali*, Ghodsi, Wong

Honors, Funding, & Awards

As a faculty

2024	**Co-PI**	Waterloo.AI Nexus of Data & AI Seed Funding	(CAD $20,000)	*Waterloo, ON*
2024	**Sole-PI**	NSERC Discovery Grant	(CAD $200,000)	*Waterloo, ON*
2024	**Sole-PI**	NSERC Discovery Grant Supplements	(CAD $12,500)	*Waterloo, ON*
2024	**Recipient**	Igor Ivkovic Teaching Excellence Award	-	*Waterloo, ON*
2023	**Recipient**	Institution Startup Fund	(CAD $170,000)	*Waterloo, ON*

As a student

2023	**Recipient**	ETH Zurich Medal for Outstanding Doctoral Performance	(CHF 2,000)	*Zürich, CH*
2021	**Recipient**	Google PhD Fellowship in Machine Learning (AI for Social Good)	(USD $210,000)	*Tübingen, DE*
2018	**Recipient**	Max Planck ETH Center for Learning Systems PhD Fellowship	-	*Tübingen, DE*
2018	**Recipient**	NSERC Postgraduate Scholarship - Doctorate (PGS-D)	(CAD $63,000)	*Waterloo, CA*
2018	**Recipient**	NSERC Canada Graduate Scholarship - Doctorate (CGS-D)	(CAD $105,000) | DECLINE	*Waterloo, CA*
2018	**Recipient**	President's Graduate Scholarship (PGS)	(CAD $35,000) | DECLINE	*Waterloo, CA*
2018	**Recipient**	David R. Cheriton Graduate Scholarship	(CAD $20,000) | DECLINE	*Waterloo, CA*
2018	**Recipient**	Alumni Gold Medal for highest standing across university	-	*Waterloo, CA*
2017	**Recipient**	Ontario Graduate Scholarship (OGS)	(CAD $15,000)	*Waterloo, CA*
2017	**Recipient**	President's Graduate Scholarship (PGS)	(CAD $10,000)	*Waterloo, CA*
2017	**Recipient**	Graduate Research Scholarship	(CAD $9,000)	*Waterloo, CA*
2016	**Recipient**	Graduate Research Scholarship	(CAD $3,000)	*Waterloo, CA*
2015	**Recipient**	Spirit of EngSci Award for outstanding community contribution	-	*Toronto, CA*
2015	**Recipient**	Dean's Honour List (7/8 semesters)	-	*Toronto, CA*
2014	**Recipient**	The Next 36 Entrepreneurship Institute	-	*Toronto, CA*
2013	**Recipient**	Shaw Design Scholarship	(CAD $5,000)	*Toronto, CA*
2012	**Recipient**	Exceptional Opportunities Award in Engineering Science	(CAD $3,000)	*Stanford, USA*
2012	**Recipient**	Peter Sands Scholarship Award in Engineering Science	(CAD $2,000)	*Toronto, CA*
2012	**Finalist**	Deloitte Consulting Competition	-	*Toronto, CA*
2012	**Finalist**	National Business & Technology Consulting Competition	-	*Toronto, CA*
2010	**Top %0.1**	Iranian National University Entrance Exam (*"Konkoor"*)	-	*Shiraz, IR*
2009	**Finalist**	Iranian Mathematics Olympiad	-	*Shiraz, IR*
2008	**3rd place**	Provincial Computer Skills Competition	-	*Shiraz, IR*
2003	**Recipient**	2 Gold, 8 Silver, & 8 Bronze medals in regional chess competitions	-	*Waterloo, CA*

Teaching

2024	**Instructor**	Intro to Machine Learning	∼ 150 students	*Waterloo, CA*
2024	**Instructor**	Intro to Machine Learning (☆ Teaching Excellence Award ☆)	∼ 100 students	*Waterloo, CA*
2023	**Co-instructor**	Socially Responsible Machine Learning: A Causal Perspective	∼ 50 attendees	*Long Beach, CA*
2022-p.	**Co-instructor**	Providing free public education on basic & advanced AI subjects on YouTube, Instagram, Substack, & Medium	∼ 25,000 students	*online*
2021	**Co-instructor**	Causethical Machine Learning	∼ 20 students	*Saarbrücken, DE*
2021	**Assistant**	Computational Intelligence Lab	∼ 150 students	*Zürich, CH*
2020	**Assistant**	Deep Learning	∼ 150 students	*Zürich, CH*
2020	**Assistant**	Causal Representation Learning	∼ 40 students	*Zürich, CH*
2018	**Head TA**	Machine Learning	∼ 250 students	*Waterloo, CA*

2017	**Assistant**	Data Types and Structures	$\sim$ 200 students	*Waterloo, CA*
2017	**Assistant**	Data Structures and Data Management	$\sim$ 200 students	*Waterloo, CA*
2017	**Assistant**	Computer Organization and Design	$\sim$ 200 students	*Waterloo, CA*
2015	**Instructor**	Taught a course on programming basics (via "Scratch"), logic, combinatorics, and computer assembly to students aged 9-12.	$\sim$ 4 students	*Shiraz, IR*
2012	**Instructor**	Tutored students on linear algebra and calculus	$\sim$ 2 students	*Toronto, CA*

Advising

Postdoc

2024	**Supervisor**	Hamdi Altahery	-	🏛 Waterloo	-

PhD

2024	**Supervisor**	Zahra Khotanlou	-	🏛 Waterloo	-
2024	**Co-Supervisor**	Mina Kebriaee	-	🏛 Waterloo	-
2023	**Mentor**	Ahmad Ehyaei	📕 C14	🏛 Tübingen	-
2022	**Mentor**	Miriam Rateike	-	🏛 Saarland	→ Google PhD Fellow 2023

Masters

2024	**Co-Supervisor**	Maryam Ghorbansabagh	-	🏛 Waterloo	-
2024	**Co-Supervisor**	Zachary Wu	-	🏛 Waterloo	-
2021	**Mentor**	Ricardo Dominguez-Olmedo	📕 C13, C15	🏛 Tübingen	→ Intl. Max Planck Research Schools
2019	**Mentor**	Alexandra Walter	-	🏛 Tübingen	→ Helmholtz Data Sci. Sch. of Health

Bachelors

2024	**Supervisor**	Mohammadreza Alavi	-	🏛 Sharif	-
2024	**Supervisor**	Abubakar Bello	-	🏛 Waterloo	→ Microsoft Inc.
2020	**Mentor**	Kiarash Mohammadi	📕 C11	🏛 Ferdowsi	→ MILA AI Institute

Invited Talks

2024	"AI for Emotion Detection"	@ BBC Radio 4	*online*
2024	"Building Bridges: Towards Trustworthy Human-AI Decision Making"	@ Toronto Machine Learning Summit	*Toronto, CA*
2024	"Building Bridges: Towards Trustworthy Human-AI Decision Making"	@ Waterloo Alumni Reunions	*Waterloo, CA*
2024	"Building Bridges: Towards Trustworthy Human-AI Decision Making"	@ Waterloo.AI Seminars	*Waterloo, CA*
2024	"Building Bridges: Towards Trustworthy Human-AI Decision Making"	@ Waterloo.AI Nexus of Data & AI Event	*Waterloo, CA*
2024	"Building Bridges: Towards Trustworthy Human-AI Decision Making"	@ Waterloo VIP Lab	*Waterloo, CA*
2024	"Building Bridges: Towards Trustworthy Human-AI Decision Making"	@ Waterloo ECE Seminars	*Waterloo, CA*
2023	"AI and the Transformation of Social Science Research"	@ University of Waterloo	*Waterloo, CA*
2023	"Towards trustworthy human-machine collaboration"	@ MIT Media Lab	*Cambridge, USA*
2023	"Towards trustworthy human-machine collaboration"	@ University of Waterloo	*Waterloo, CA*
2023	"Towards trustworthy human-machine collaboration"	@ University of Alberta	*Alberta, CA*
2023	"Towards trustworthy human-machine collaboration"	@ Carnegie Mellon Uni.	*Doha, QA*
2023	"Towards trustworthy human-machine collaboration"	@ New York University	*Abu Dhabi, UAE*
2022	"Improving search via type- & IO-based neurally-guided program synthesis"	@ Google DeepMind	*London, UK*
2022	"Toward Invariant Explanations"	@ Google Brain (CAM)	*online*
2022	"Algorithmic recourse: from theory to practice"	@ Google Brain (LUMA)	*online*
2022	"Algorithmic recourse: from theory to practice"	@ Google DeepMind	*London, UK*
2022	"Algorithmic recourse: from theory to practice"	@ MILA	*online*
2022	"Algorithmic recourse: from theory to practice"	@ IMS Annual Meeting	*London, UK*
2021	"Algorithmic recourse: from theory to practice"	@ Harvard University	*online*
2021	"Algorithmic recourse: from theory to practice"	@ NEC Europe Labs	*online*
2021	"Algorithmic recourse: from theory to practice"	@ Cyber Valley Health	*online*

2020	"Algorithmic recourse: from theory to practice"	@ ETH IML Seminars		*online*
2020	"Algorithmic recourse: from theory to practice"	@ UCL Causality Group		*online*
2021	"Algorithmic Recourse: from Counterfactual Explanations to Interventions"	@ ACM FAccT'21		*online*
2020	"Algorithmic recourse under imperfect causal knowledge"	@ NeurIPS'20		*online*
2020	"Model-Agnostic Counterfactual Explanations for Consequential Decisions"	@ AISTATS'20		*online*
2019	"Model-Agnostic Counterfactual Explanations for Consequential Decisions"	@ MPI-IS Retreat		*Blaubeuren, DE*
2016	"The decision to join, and the decision to leave Facebook"	@ Vis. & Img. Proc. Lab		*Waterloo, CA*
2015	"From Shahed to Broadway"	@ Alma mater highsch.		*Shiraz, IR*
2015	"The Perfect Human" ("Al Ensan Al Kamel" - book excerpt)	@ UofT Qur'an Group		*Toronto, CA*
2015	"100 percent of the shots you don't take are a miss!"	@ Galbraith Society		*Toronto, CA*

Reviewing

2024-p.	**Reviewer**	Journal of Artificial Intelligence	≡: *Art. Int.*	∣
2021-p.	**Reviewer**	Journal of Machine Learning Research	≡: *JMLR*	∣
2021-p.	**Prog. Comm.**	ACM Fairness, Accountability, and Transparency	≡: *ACM FAccT*	∣
2021-p.	**Prog. Comm.**	International Conference on Learning Representations	≡: *ICLR*	∥
2021-p.	**Prog. Comm.**	International Conference on Machine Learning	≡: *ICML*	∣
2021-p.	**Prog. Comm.**	European Conference on Machine Learning	≡: *ECML*	∣
2020-p.	**Prog. Comm.**	Artificial Intelligence and Statistics	≡: *AISTATS*	∥∣
2019-p.	**Prog. Comm.**	Neural Information Processing Systems	≡: *NeurIPS*	∥∣

Other Service

2024-p.	**Comm. Mbr.**	PhD Defence	∣	*Waterloo, CA*
2024-p.	**Chair**	PhD Comprehensive Exam	∣	*Waterloo, CA*
2024-p.	**Comm. Mbr.**	PhD Comprehensive Proposal Exam	∣	*Waterloo, CA*
2024-p.	**Comm. Mbr.**	PhD Comprehensive Background Exam	∥∥	*Waterloo, CA*
2024	**Comm. Mbr.**	Vector Institute Scholarship in AI Review Committee	-	*Toronto, CA*
2024	**Comm. Mbr.**	ICML Workshops Proposal Review Committee	-	*online*
2023	**Co-organizer**	ICML Workshop on Counterfactuals in Minds & Machines	∼ 50 attendees	*Hawaii, USA*
2021	**Volunteer**	Max Planck ETH Center for Learning Systems Retreat	∼ 50 attendees	*online*
2021	**Co-organizer**	ELLIS Workshop on Causethical ML	∼ 50 attendees	*online*
2021	**Co-organizer**	ICML Workshop on Algorithmic Recourse	∼ 50 attendees	*online*
2020	**Co-organizer**	NeurIPS Symposium on Algorithmic Recourse	∼ 100 attendees	*online*
2020	**Note-taker**	CIFAR Causal Inference Program Meeting	∼ 50 attendees	*online*
2020	**Volunteer**	European Conference on Machine Learning (ECML)	-	*online*
2020	**Volunteer**	Machine Learning Summer School (MLSS)	∼ 150 attendees	*online*
2016	**Volunteer**	Conference on Vision and Imaging Systems (CVIS)	∼ 150 attendees	*Waterloo, CA*
2016	**Volunteer**	Conference on Vision and Imaging Systems (CVIS)	∼ 150 attendees	*Waterloo, CA*
2013	**Logistics Dir.**	Engineering Science Educational Conference (ESEC)	∼ 600 attendees	*Toronto, CA*
2012	**Volunteer**	Engineering Science Educational Conference (ESEC)	∼ 600 attendees	*Toronto, CA*
2012	**Finance Dir.**	Cut for Cancer - Shave the Prof. (Students Fighting Cancer)	∼ 350 attendees	*Toronto, CA*