

Ramaravind Kommiya Mothilal

PhD Candidate, University of Toronto
Data Sciences Institute (DSI) Graduate Fellow

[✉ ramaravind.com](http://ramaravind.com) [@ ram.mothilal@mail.utoronto.ca](mailto:ram.mothilal@mail.utoronto.ca) [GitHub](https://github.com/ramaravindkm) [Google Scholar](https://scholar.google.com/citations?user=QWzJyQAAAAJ&hl=en) [in ramaravindkm](https://www.linkedin.com/in/ramaravindkm)

Research Objective

My overall research goal is to support the safe and responsible use of AI by (a) analyzing the norms, tensions, and subtle technical assumptions that underlie AI system development, and (b) developing software and frameworks that translate theoretical and qualitative analyses into practical mechanisms for the responsible deployment of AI in real-world settings.

Education

Present	University of Toronto PhD in Information	Toronto, Canada
Sep 2022	> Advisors: Prof. Syed Ishtiaque Ahmed and Prof. Shion Guha > Topic: Interpreting how AI practitioners reason through ambiguities in responsible AI, and complementarily, evaluating and improving the reasoning-like behavior of LLMs under ambiguous safety-related contexts, via a mixed-method approach (empirical qualitative studies; literature reviews; quantitative evaluation frameworks)	
Feb 2018	Katholieke Universiteit Leuven (KU Leuven)	Leuven, Belgium
Sep 2016	MSc in Artificial Intelligence (<i>Engineering and Computer Science Specialization</i>) > Percentage: 76.81% (Magna Cum Laude) > Thesis: Statistical Modeling of Students' Performance in an Open-Admission Bachelor Program in Flanders - The need to discriminate between explanatory and predictive modeling [🔗]	
May 2015	National Institute of Technology (NIT) Trichy	Tiruchirappalli, India
July 2011	Bachelor in Instrumentation and Control Engineering > CGPA: 8.1/10 (First Class)	

Experience

Present	University of Toronto Family Housing	Toronto, Canada
Aug 2025	Residence Advisor > Conducted a hands-on beginner-friendly series to help students, professionals, and lifelong learners build key skills in online writing with LaTeX/Overleaf, managing literary sources, coding with Python, and using AI tools like ChatGPT responsibly.	
Jun 2022	Everwell Health Solutions	Bangalore, India
Mar 2021	Data Scientist / Mentor: Dr. Bill Thies > Designed a systematic data hygiene framework to anticipate and remove sources of data bias in downstream ML tasks. > Developed indicators to evaluate and improve engagement of health providers to facilitate patient adherence to medications. > Built various dashboards with Power BI and Apache Superset for different stakeholders, including USAID and Central TB Division—Government of India.	
Feb 2021	Microsoft Research Center for Societal Impact through Cloud & AI [🔗]	Bangalore, India
Jan 2019	AI Center Fellow / Advisors: Dr. Amit Sharma, Dr. Joyojeet Pal > Developed a general optimization framework—DiCE(Diverse Counterfactual Explanations)—to generate and evaluate a diverse set of hypothetical examples that show people how to obtain a different prediction. [FAccT'20] [🔗] [🔗] > Studied the complementarities of Feature Attribution and Counterfactual explanation methods to demonstrate the value of considering multiple kinds of model explanations. [AIES'21] > Studied how occupying (un)favourable positions in Indian politicians' social networks leads to inequality in terms of both information received and propagation of new information [CSCW'21]	
Dec 2018	Microsoft Research Technology for Emerging Markets [🔗]	Bangalore, India
July 2018	Research Intern / Advisors: Dr. Amit Sharma	
Jul 2016	Royal Bank of Scotland (RBS) Data Archival Team	Chennai, India
Jul 2015	Analyst	

- [P.2] **Reasoning About Reasoning: Towards Informed and Reflective Use of LLM Reasoning in HCI** [🔗]
Ramaravind Kommiya Mothilal, Sally Zhang, Syed Ishtiaque Ahmed, and Shion Guha
- [P.1] **Human-Aligned Faithfulness in Toxicity Explanations of LLMs** [🔗]
Ramaravind Kommiya Mothilal, Joanna Roy, Syed Ishtiaque Ahmed, and Shion Guha
- [C.13] **Talking About the Assumption in the Room** [🔗]
Ramaravind Kommiya Mothilal, Faisal M Lalani, Syed Ishtiaque Ahmed, Shion Guha, and Sharifa Sultana
2025 CHI Conference on Human Factors in Computing Systems [CHI'25]
- [C.12] **Towards a Non-Ideal Methodological Framework for Responsible ML** [🔗]
Ramaravind Kommiya Mothilal, Shion Guha, Syed Ishtiaque Ahmed
2024 CHI Conference on Human Factors in Computing Systems [CHI'24]
- [C.10] **Caste Capital on Twitter: A Formal Network Analysis of Caste Relations among Indian Politicians** [🔗]
Ramaravind Kommiya Mothilal*, Palashi Vaghela*, Joyojeet Pal (* = Equal Contribution)
Proceedings of the ACM on Human-Computer Interaction (CSCW) [CSCW'22]
- [C.9] **Voting with the stars: Analyzing Partisan Engagement between Celebrities and Politicians in India** [🔗]
Ramaravind Kommiya Mothilal, Dibyendu Mishra, Sachita Nishal, Faisal M. Lalani, Joyojeet Pal
Proceedings of the ACM on Human-Computer Interaction (CSCW) [CSCW'22]
- [C.8] **Towards unifying feature attribution and counterfactual explanations: Different means to the same end** [🔗]
Ramaravind Kommiya Mothilal, Divyat Mahajan, Chenhao Tan, Amit Sharma
Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, 652-663 [AIES'21]
- [C.7] **Birds of a Caste-How Caste Hierarchies Manifest in Retweet Behavior of Indian Politicians** [🔗]
Palashi Vaghela*, Ramaravind Kommiya Mothilal*, Joyojeet Pal (* = Equal Contribution)
Proceedings of the ACM on Human-Computer Interaction [CSCW'21]
- [C.6] **Using Mobile Airtime Credits to Incentivize Learning, Sharing and Survey Response: Experiences from the Field** [🔗]
Devansh Mehta, Ramaravind Kommiya Mothilal, Alok Sharma, William Thies, Amit Sharma
Proceedings of the 3rd ACM SIGCAS Conference on Computing and Sustainable Societies [COMPASS'20]
- [C.5] **Topical Focus of Political Campaigns and its Impact: Findings from Politicians' Hashtag Use during the 2019 Indian Elections** [🔗]
Anmol Panda, Ramaravind Kommiya Mothilal, Monojit Choudhury, Kalika Bali, Joyojeet Pal
Proceedings of the ACM on Human-Computer Interaction [CSCW1'20]
- [C.4] **Learnings from Technological Interventions in a Low Resource Language: A Case-Study on Gondi** [🔗]
Sebastin Santy, Devansh Mehta, Ramaravind Kommiya Mothilal, Brij Mohan Lal Srivastava, Alok Sharma, Anurag Shukla, Vishnu Prasad, Venkanna U, Amit Sharma and Kalika Bali
International Conference on Language Resources and Evaluation [Invited Submission @ LRE Journal] [LREC'20]
- [C.3] **Explaining machine learning classifiers through diverse counterfactual explanations** [🔗]
Ramaravind Kommiya Mothilal, Amit Sharma, Chenhao Tan
Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency [FAccT'20]
- [C.2] **Optimizing peer referrals for public awareness using contextual bandits** [🔗]
Ramaravind Kommiya Mothilal, Amulya Yadav, Amit Sharman
Proceedings of the 2nd ACM SIGCAS Conference on Computing and Sustainable Societies [COMPASS'19]
- [C.1] **Predicting first-year engineering student success: from traditional statistics to machine learning** [🔗]
Ramaravind Kommiya Mothilal, Tom Broos, Maarten Pinxten, Tinne De Laet
Proceedings of the 46th SEFI Annual Conference 2018 [SEFI'18]
- [W.2] **Community-Driven Data Practices for Advancing Ethical and Equitable AI for Low-Resource Language** [🔗]
Charles Nimo, Shuheng Liu, Amy Z Chen, Ramaravind Kommiya Mothilal, Michael Best
Proceedings of the ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies [COMPASS'25]
- [W.1] **Towards Unifying Feature Attribution and Counterfactual Explanations**
Ramaravind Kommiya Mothilal, Divyat Mahajan, Chenhao Tan, Amit Sharma
Workshop on Responsible AI, ICLR-21 [Spotlight Contributed Talk] [RAI@ICLR'21]
- [D.3] **Facilitating Media Distribution with Monetary Incentives** [🔗]
Devansh Mehta, Alok Sharma, Ramaravind Kommiya Mothilal, Anurag Shukla, Vishnu Prasad, William Thies, Venkanna U, Colin Scott, Amit Sharma
Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems [CHI'20]

- [D.2] **Learnings from an Ongoing Deployment of an IVR-based Platform for Voter Awarenessss** [🔗]
Ramaravind Kommiya Mothilal, Devansh Mehta, Alok Sharma, William Thies, Amit Sharma
Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing [a non-archival version won Outstanding Poster Award at COMPASS 2019] [CSCW'19]
- [D.1] **The Appeal of Influencers to the Social Media Outreach of Indian Politicians** [🔗]
Faisal M Lalani, Ramaravind Kommiya Mothilal, Joyojeet Pal
Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing [COMPASS'19]

Computing Skills

- > **Operating System:** Windows, Linux
- > **Database/Server:** MySQL, PostgreSQL, SQL Server
- > **Language:** Python, R, C/C++, Java, HTML, CSS, JavaScript, PHP, SQL, Hadoop, Spark
- > **Machine Learning Frameworks:** PyTorch, TensorFlow, Hugging Face Transformers
- > **Visualization Software/Tools:** Power BI, Apache Superset, Tableau

Software and Open-Source Contribution

- > **HAF: Human-Aligned Faithfulness in Toxicity Explanations of LLMs** : University of Toronto - [commits](#), [repository](#)
- > **DiCE: Diverse Counterfactual Explanations for Machine Learning** : Microsoft Research India - [commits](#), [repository](#)

Awards

- > **Data Sciences Institute (DSI) Graduate Fellowship - C\$75,000 - 2024-26**
- > **Schwartz Reisman Institute for Technology and Society (SRI) Graduate Fellowship - C\$7,500 - 2023-24**

Academic Service and Volunteering

Reviewer	CHI'26, CHI'25, FAccT'25, CSCW'25, CHI'24, DIS'24, SatML'24 ICLR-RAI'21, TOCHI'22, GHTC'22
Reviewer & AC	CHI'24 Late Breaking Works
Reviewer & PC	COMPASS'25, COMPASS'22, COMPASS'21, XKDD'21
Co-Organizer	COMPASS'25 Late Breaking Work Chair
Mentor	2024 Responsible Tech Mentorship Program by All Tech Is Human
Co-Supervisor	UofT CS and EngSci undergraduate students: Joanna Roy and Jade Clement
Section Leader	<i>Code in Place</i> by Stanford University

Teaching

Teaching Assistant, UofT	Speculative Design (INF 2260), Experimental Design for Data Science (INF 2178), Computational Thinking (INF 1339), Programming for Data Science (INF 1340), Data Analytics (INF 2190), Data Modeling and Database Design (INF 1343), Human Values in Data Science (INF 2210), Computers and Society (CSC 300), Statistics for Data Science (INF 1344)
Workshop Presenter	Data Analysis with SQL and MySQL, iSkills Workshop at Faculty of Information, UofT (Fall 2022, Winter 2023, Fall 2023, Winter 2024)

Mentions and Coverages

- > “Dalit MPs less likely to be heard on Twitter than upper castes” | [The Times of India](#) Oct 2020
- > “Birds of Brahmanical Feather Flock Together” | [The Wire](#) Aug 2020
- > “DiCE: Diverse Counterfactual Explanations for Hotel Cancellations” | [Towards Data Science](#) Jul 2020
- > “DiCE -ML models with counterfactual explanations for the sunk Titanic-” | [Analytics Vidhya](#) Mar 2020
- > “Open-source library provides explanation for machine learning through diverse counterfactuals” | [Microsoft Research Blog](#) Jan 2020
- > “Elections 2019: What tweet for tat tells you about Rahul Gandhi” | [LiveMint](#) Apr 2019