Aniket **Vashishtha**

aniketvashishtha.github.io @ aniketv2@illinois.edu Google Scholar +1-447-902-6632

Education

May 2026 | **University of Illinois Urbana-Champaign** — **Illinois, USA**
Aug 2024 | MS in Computer Science *(Thesis Track)*, GPA: 4.0/4.0

Aug 2022 | **Guru Gobind Singh Indraprastha University** — **New Delhi, India**
Aug 2018 | B.Tech., Information Technology, GPA: 8.5/10.0

Experience

Dec 2025 – Aug 2025 | **University of Illinois Urbana-Champaign** [🌐] — **Illinois, USA**
Research Assistant Advisor: Prof. Hao Peng, Prof. Chenhao Tan, Prof. Jiawei Han
- Developed a code-based framework to evaluate and improve LLMs' counterfactual reasoning via Reinforcement Learning (RL), demonstrating strong generalization across out-of-domain settings.
- Building a framework for extracting causal graph structures via RAG for multi-hop reasoning.

Aug 2024 – Apr 2023 | **Microsoft Research** [🌐] — **Bangalore, India**
Pre-Doctoral Research Fellow - Advisor: Dr. Amit Sharma, Prof. Vineet Balasubramanian
- Developed methods to leverage imperfect experts (like humans and LLMs) for accurate discovery of causal relationships and effect estimation. Work accepted at **ICLR'25** conference.
- Proposed an axiomatic framework for causal reasoning and trained a 67M model from scratch that outperformed Gemini Pro and GPT-4 on applying causal axioms. Work accepted at **ICML'25**.

Aug 2022
AI Center Fellow Advisor: Sameer Segal
Led a multifaceted effort merging HCI research, engineering, design, and psychology to build a workplace mental health application currently deployed at Microsoft Research India.

Jan 2022
Research Intern - Advisors: Prof. Monojit Choudhury, Dr. Sunayana Sitaram
- Led a Responsible AI initiative assessing bias in LLM-based auto-suggestions, created checklists to measure social bias and evaluate effects of harm-mitigation strategies.Work accepted at **EACL'23**
- Proposed improved bias metrics for multilingual language models and introduced evaluation and debiasing methods for Indian languages to enhance inclusivity. Work accepted at **ACL'23**.

Jan 2022 – Mar 2021 | **Inria** [🌐] — **Remote/Paris, France**
Research Intern - Advisors: Dr. Adrien Coulet, Dr. Joel Legrand
- Worked on the identification of discontiguous entities using segmental hypergraph and dependency graphs on Pharmacogenomics corpora.

Apr 2021 – Sep 2020 | **IIIT Delhi - TavLab Research Group** [🌐] — **New Delhi, India**
Research Intern - Advisor: Prof. Tavpritesh Sethi
- Optimized vaccine allocation for Indian states via RL, contextual bandits and compartmental models. Also conducted a spacio-temporal analysis of the vaccine infodemic on social media.

Selected Publications W = Workshop, C=Conference, J=Journal (* = Equal Contribution)

[W.1] **Executable Counterfactuals: Improving LLMs' Causal Reasoning Through Code** [🔗]
Aniket Vashishtha*, Qirun Dai*, Hongyuan Mei, Amit Sharma, Chenhao Tan, Hao Peng
Foundations of Reasoning in Language Models Workshop@NeurIPS'25 [**FoRLM@Neurips'25**]

[C.1] **Teaching Transformers Causal Reasoning through Axiomatic Training** [🔗]
Aniket Vashishtha, Abhinav Kumar, Atharva Pandey, Abbavaram Gowtham Reddy, Kabir Ahuja, Vineeth N. Balasubramanian, Amit Sharma
Forty-Second International Conference on Machine Learning [**ICML'25**]

[C.2] **Causal Order: The Key to Leveraging Imperfect Experts in Causal Inference** [🔗]
Aniket Vashishtha, Abbavaram Gowtham Reddy, Abhinav Kumar, Saketh Bachu, Vineeth N. Balasubramanian, Amit Sharma
The Thirteenth International Conference on Learning Representations [**ICLR'25**]

[W.2] **Realizing LLMs' Causal Potential Requires Science-Grounded, Novel Benchmarks** [🔗]
Ashutosh Srivastava, Lokesh Nagalapatti, Gautam Jajoo, Aniket Vashishtha, Parameswari Krishnamurthy, Amit Sharma
CauScien: Uncovering Causality in Science [**CauseScien@NeurIPS'25**]

[C.3] **On Evaluating and Mitigating Gender Biases in Multilingual Settings** [🔗]
Aniket Vashishtha*, Kabir Ahuja*, Sunayana Sitaram
Annual Conference of the Association for Computational Linguistics [**ACL'23 Findings**]

[C.4] **Performance and Risk Trade-offs for Multi-word Text Prediction at Scale** [🔗]
Aniket Vashishtha, S Sai Krishna Prasad, Payal Bajaj, Vishrav Chaudhary, Kate Cook, Sandipan Dandapat, Sunayana Sitaram, Monojit Choudhury
European Chapter of the Association for Computational Linguistics [**EACL'23 Findings**]

[J.1] **Mining Trends of COVID-19 Vaccine Beliefs on Twitter With Lexical Embeddings: Longitudinal Observational Study** [🔗]
Harshita Chopra*, Aniket Vashishtha*, Ridam Pal, Ashima, Ananya Tyagi, Tavpritesh Sethi
Journal of Medical Internet Research Infodemiology [**JMIR Infodemiology'23**]

[J.2] **VacSIM: Learning effective strategies for COVID-19 vaccine distribution using reinforcement learning** [🔗]
Raghav Awasthi, Keerat Kaur Guliani, Saif Ahmad Khan, Aniket Vashishtha, Mehrab Singh Gill, Arshita Bhatt, Aditya Nagori, Aniket Gupta, Ponnurangam Kumaraguru, Tavpritesh Sethi
Intelligence Based Medicine Journal [**IBM'22**]

Selected Research Projects

Executable Counterfactuals: Improving LLM's Causal Reasoning through Code Jan'25 - Sept'25
Advisor: Prof. Hao Peng, Prof. Chenhao Tan, Dr. Amit Sharma

- Proposed *Executable Counterfactuals*, a code-based framework for evaluation counterfactual (CF) reasoning abilities of LLMs. Also built GSM-style counterfactual math problems using dependency graphs to test cross-domain generalization of CF reasoning skills.
- Trained Qwen models (1.5B–7B) via SFT and RL on simple if-else based code CF tasks, and evaluated transfer to harder OOD code logic. Found that SFT improves in-domain accuracy, while RL with outcome-based rewards (GRPO) induces transferable CF reasoning skills that generalize to more complex code and math problems.
- Work currently **under review at ICLR'26**, and accepted at **FoRLM workshop, NeurIPS'25**

Teaching Transformers Causal Reasoning through Axiomatic Training Feb'24 - Jan'25
Advisors: Dr. Amit Sharma, Prof. Vineeth Balasubramanian

- Introduced an *Axiomatic Framework*, a new training paradigm for language models, enabling a 67M-parameter model to outperform GPT-4 on applying causal axioms such as *transitivity* and *d-separation* in complex graphs.
- Explored transformer architectural adjustments and axiomatic fine-tuning of LLMs to improve generalization to unseen causal structures and inferring causal relationships from natural-language correlational statements.
- Work published at **ICML'25** and also covered by multiple research forums and blog posts [1],[2],[3]

Causal Order: The Key to Leveraging Imperfect Experts in Causal Inference Apr'23 - Aug'24
Advisors: Dr. Amit Sharma, Prof. Vineeth Balasubramanian

- Studied imperfect experts (humans and LLMs) for real-world causal discovery, showing that *causal order* is both easy to extract and robust, yet critical for downstream performance.
- Proposed an optimal method for extracting domain knowledge via a triplet-based querying framework and integrating it as priors into statistical discovery algorithms yielding strong gains on real-world healthcare datasets.
- Work published at **ICLR'25**, presented as an **oral presentation** at the *CaLM@NeurIPS'24* (nominated for Best Paper award) & *LLM-CP@AAAI'24* workshops, and slated to be integrated with **DoWhy/PyWhy**

Performance and Risk Trade-offs for Multi-word Text Prediction at Scale Jan'22 - Oct'22
Advisor: Prof. Monojit Choudhury, Dr. Sunayana Sitaram

- Assessed toxicity detection methods using our proposed CheckList based dataset for improving fairness of text prediction systems, targeting various harm levels across different demographic groups.
- Led an annotation study with diverse participants to build a benchmark of sensitive attributes for evaluating how LLMs produce harmful auto-suggestions and their contextual severity.
- Work published at **EACL'23** Findings, with the toxicity benchmark released publicly.

Talks

"Executable Counterfactuals: Improving LLMs Causal Reasoning through Code "[🌐] February 2026

› Invited to give a talk at the statistics seminar of Indiana University Indianapolis.

"Teaching Transformers Causal Reasoning through Axiomatic Training "[🌐] November 2024

› Gave a talk at the Causal Data Science Summit (CDSM'25), a prestigious event by Maastricht University and Copenhagen Business School for research in causality and economics.

"Causal Inference using LLM-Guided Discovery" [🌐] February 2024

› Gave a talk at *LLM-CP workshop* for our work accepted as an **Oral** presentation at **AAAI'24**.

"Empowering Education, Enhancing Experiences and Beyond with LLMs" [🌐] September 2023

› Led a 2-day national workshop on LLMs and practical rural education strategies for faculty and students from non computer science disciplines in Kerala.

"How To Break Into Data Science and AI Research"[🌐] March 2022

› Delivered a talk on how to pick up Data Science and research in AI to 200+ undergraduate students

Recognition and Awards

Recipient of Thinking Machines' Research Grant [🌐]
Received USD $5,000 Tinker Research Grant from Thinking Machines Lab for research on causality and LLMs.

Interview on Causality and LLM Research Featured by Neptune.ai [📹]
My research on *Axiomatic Training* for improving causal reasoning of LLMs was covered by Neptune.ai at ICML'25

Research featured in Causal Bandits Newsletter and Podcast [🌐][📹]
Research on causality and LLMs was covered in the newsletter and podcast of the Causal Bandits series

Awarded Travel Grant Award of 1500 USD for ICLR'25 [🌐]
Received travel grant for 1500 USD to present my research at ICLR'25 in Singapore.

Spotlight Presentations at NeurIPS'24 and AAAI'24 workshops [📹 📹]
Research on causality and LLMs was presented in spotlight talks at CaLM@NeurIPS'24 and LLM-CP@AAAI'24.

Recipient of the JN Tata Scholarship [🌐]
Recipient of the prestigious JN Tata Scholarship, a competitive merit-based scholarship loan awarded to a select number of students across India for pursuing higher education.

Winner of Microsoft Turing's Large Scale Models for Inclusion Hackathon Challenge, 2022 [🌐]
Developed an Inclusivity Toolkit to diagnose the biases of language models across various dimensions by bringing together numerous bias detection methods in the literature

High Commendation Award for Trinity Challenge, 2021
Contributed to the submission, *"Privacy-preserving Crowdsourcing for Citizen Engagement in Pandemics*", which was recognized as a Highly Commended Solution at The Trinity Challenge 2021.

Spotlight Presentation at MIT's Conference 'Vaccines for All', 2020 [🌐] [📹]
Research on building RL based frameworks for vaccine allocation got featured as a part of the spotlight presentation

Leadership and Volunteering Roles

Causal NLP Reading Group, MSR India *Founding Member and Coordinator* Apr'23 - Aug'24

Started a weekly reading group on causality & LLMs, hosting speakers from University of Tornto, IIT Bombay

PathCheck Foundation (MIT) *Data Science Researcher* Feb'21 - May'21

Developed data-driven COVID-19 response strategies using data driven strategies to support public health planning

Red Dot Foundation - Safecity *Data Analytics Volunteer* Apr'20 - Dec'20

Led data-driven initiatives to combat harassment against women and support the LGBTQ+ community in Indian cities through crowdsourced crime reporting, chatbot development, inclusive education.

Academic Service

Reviewer	ICLR'26, COLM'25, ACL'25, Re-Align Workshop @ICLR'25, Reasoning & Planning for LLMs@ICLR'25, Behaviour ML @NeurIPS'24, 4th Workshop on NLP for Indigenous Languages of the Americas'24, LREC'22
Teaching	TA at UIUC for CS 307: Modeling & Learning in Data Science; CS 277: DSA for Data Science