



# Aniket Vashishtha

 [aniketvashishtha.github.io](https://github.com/aniketvashishtha) @ [aniketv2@illinois.edu](mailto:aniketv2@illinois.edu)  [Google Scholar](#) +1-447-902-6632

## EDUCATION




May 2026 Aug 2024	<b>University of Illinois Urbana-Champaign</b> MS in Computer Science ( <i>Thesis Track</i> ), GPA: 4.0/4.0	Illinois, USA
Aug 2022 Aug 2018	<b>Guru Gobind Singh Indraprastha University</b> B.Tech., Information Technology, GPA: 8.5/10.0	New Delhi, India

## EXPERIENCE

Aug 2024 Aug 2022	<b>University of Illinois Urbana-Champaign</b>  <i>Research Assistant Advisor: Prof. Hao Peng, Prof. Chenhao Tan, Prof. Jiawei Han</i> › Developed a code-based framework to evaluate and improve LLM counterfactual reasoning via RL, demonstrating strong generalization across out-of-domain settings. › Building a framework for extracting causal graph structures via RAG for multi-hop reasoning.	Bangalore, India
Aug 2024 Apr 2023	<b>Microsoft Research</b>  <i>Pre-Doctoral Research Fellow - Advisor: Dr. Amit Sharma, Prof. Vineet Balasubramanian</i> › Developed methods to leverage imperfect experts (LLMs) for accurate causal discovery and effect estimation. Work accepted at <b>ICLR'25</b> conference. › Proposed an axiomatic framework for causal reasoning and trained a 67M model from scratch that outperformed Gemini Pro and GPT-4 on complex causal graphs. Accepted at <b>ICML'25</b> .	Bangalore, India
Aug 2022	<i>AI Center Fellow Advisor: Sameer Segal</i> Led a multifaceted effort merging HCI research, software development, design, and psychology to build a workplace mental health application currently deployed at Microsoft Research India.	
Jan 2022	<i>Research Intern - Advisors: Prof. Monojit Choudhury, Dr. Sunayana Sitaram</i> › Led a Responsible AI initiative assessing bias in LLM-based text models, creating checklists to measure social bias and evaluate effects of harm-mitigation strategies. Work accepted in <b>EACL'23</b> › Proposed improved bias metrics for multilingual LMs and introduced evaluation and debiasing methods for Indian languages to enhance inclusivity. Accepted at <b>ACL'23</b> .	
Jan 2022 Mar 2021	<b>Inria</b> <i>Research Intern - Advisors: Dr. Adrien Coulet, Dr. Joel Legrand</i> Worked on the identification of discontinuous entities using segmental hypergraph and dependency graphs on Pharmacogenomics corpora.	Remote/Paris, France
Apr 2021 Sep 2020	<b>IIIT Delhi - TavLab Research Group</b>  <i>Research Intern - Advisor: Prof. Tavpritesh Sethi</i> Optimized vaccine allocation for Indian states via Reinforcement Learning and compartmental models. Also conducted a spacio-temporal analysis of the vaccine infodemic using NLP.	New Delhi, India

## 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@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]

- [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]

## SELECT RESEARCH PROJECTS

---

### Executable Counterfactuals: Improving LLM’s Causal Reasoning through Code JAN’22 - OCT’22

*Advisor: Prof. Hao Peng, Prof. Chenhao Tan, Dr. Amit Sharma*

- > Proposed Executable Counterfactuals, a code-based framework for evaluation counterfactual reasoning abilities of LLMs. Also built GSM-style counterfactual (CF) 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 CF if-else tasks with one hidden variable, evaluating transfer to harder OOD logic for CF reasoning. While SFT improves in-domain accuracy, RL with outcome-based rewards (GRPO) induces transferable CF skills that generalize across 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 - AUG’24

*Advisors: Dr. Amit Sharma, Prof. Vineeth Balasubramanian*

- > Introduced *Axiomatic Framework*, a new training paradigm for language models. Our 67M-parameter model, trained on simple causal chains, outperforms billion-scale LLMs and rivals GPT-4 in application of transitivity and d-separation in complex graphs for inferring causal relationships.
- > Architectural adjustments in transformers are explored to improve their generalization to unseen causal structures, also extending to complex tasks like inferring causality from correlational statements.
- > Work accepted at **ICML’25**, workshops at Neurips and ICLR, and also covered by multiple outlets [1],[2]

### Causal Order: The Key to Leveraging Imperfect Experts in Causal Inference APR’23 - AUG’24

*Advisors: Dr. Amit Sharma, Prof. Vineeth Balasubramanian*

- > Evaluated imperfect experts such as humans and LLMs in identifying real-world causal relationships, highlighting the importance of causal order in downstream tasks and showing LLMs’ potential for accurate order prediction.
- > Incorporated innovative probing strategies for precise causal order estimation, surpassing limitations of prior pairwise methods. Integrated In-Context Learning and Chain-of-Thought to optimize LLMs for causal reasoning.
- > Leveraged expert outputs as priors to enhance discovery algorithms like constraint-based and score-based methods. Our LLM pipelines yielded substantial improvements in causal order inference compared to baselines.
- > Work accepted at **ICLR 2025**, presented at the *CaLM Workshop* as an **oral presentation** at **CaLM@Neurips 2024**, and will also be integrated with **DoWhy** & **PyWhy** package for causal inference.

### 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.
- > Created a benchmark dataset encompassing various social facets (sexual orientation, nationality, etc.) to evaluate LLM’s performance in predicting toxic content, and assessing prediction severity based on context.
- > Work accepted at **EACL’23 Findings**

## TALKS

---

- “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 academics and industry professionals focused on research in causality and economics.
- “Causal Inference using LLM-Guided Discovery”** [🌐] November 2024  
‣ Gave a talk at *LLM-CP workshop* for our work accepted as an **Oral** presentation at **AAAI’24**.
- “Unlocking the Potential of Language Models: Empowering Education, Enhancing Experiences and Beyond”** [🌐] September 2023  
‣ Led a 2-day workshop in Kerala on the potential of LLMs and practical strategies for rural education, tailored for non-technical participants, including professors and students from fields like Botany and Physics.
- “On Evaluating and Mitigating Gender Biases in Multilingual Settings”** June 2023  
‣ Delivered a talk on our ACL’23 paper in the reading group of Speech & NLP Group, MSR India
- “How To Break Into Data Science and AI Research”** [🌐] March 2022  
‣ Delivered a talk on how to pickup Data Science and research in AI to 200+ undergraduate students

## HONOURS AND AWARDS

---

- Research Interview Featured by Neptune.ai** [🎥]  
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 Causal Bandits series
- Receipient 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.
- Recieved Travel Grant Award of 1500 USD for ICLR’25** [🌐]  
Recieved travel grant for 1500 USD to present my research at ICLR’25 in Singapore.
- 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
- Spotlight Presentation at MIT’s Conference ‘Vaccines for All’, 2020** [🌐] [🎥]  
Research on building RL based framework for vaccine allocation got featured as a part of the spotlight presentation
- High Commendation Award for Trinity Challenge, 2021**  
Our submission, “*Privacy-preserving Crowdsourcing for Citizen Engagement in Pandemics*”, was recognized as a Highly Commended Solution at The Trinity Challenge.

## LEADERSHIP AND VOLUNTEERING ROLES

---

- Causal NLP Reading Group, MSR India** *Founding Member and Coordinator* APR’23 - AUG’24  
Started a weekly reading group to discuss research in causal machine learning, LLMs, and related areas. Hosted speakers from University of Toronto, IIT Bombay, and other insitutions.
- PathCheck Foundation (MIT)** *Data Science Researcher* FEB’21 - MAY’21  
Contributed to COVID-19 solutions at PathCheck Foundation, an MIT spinoff, by developing impactful strategies for mitigating the effects of the pandemic.
- 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