

# Chahat Raj

## PhD in Computer Science | George Mason University

🌐 [chahatraj.github.io](https://chahatraj.github.io) @ [craj@gmu.edu](mailto:craj@gmu.edu) 🌐 [github.com/chahatraj](https://github.com/chahatraj) 🎓 Google Scholar

### Education

<b>Present</b> <b>Aug 2022</b>	<b>George Mason University</b> Ph.D. in Computer Science Advisors: <a href="#">Ziwei Zhu</a> , <a href="#">Antonios Anastasopoulos</a>	<b>Fairfax, USA</b>
<b>Aug 2021</b> <b>Aug 2019</b>	<b>Delhi Technological University (DTU)</b> Masters in Information Systems (Research Track) Advisor: <a href="#">Priyanka Meel</a>	<b>Delhi, India</b>
<b>May 2019</b> <b>Aug 2015</b>	<b>Indira Gandhi Delhi Technical University for Women (IGDTUW)</b> Bachelors in Computer Science & Engineering	<b>Delhi, India</b>

### Experience

<b>Sep 2024</b> <b>May 2024</b>	<b>University of Washington   The Information School</b> <i>Researcher</i>   Advisor: <a href="#">Aylin Caliskan</a> <u>Project</u> : Social Biases in Visual Question Answering in Generative AI.	<b>Seattle, USA</b>
<b>Aug 2024</b> <b>May 2024</b>	<b>George Mason University   School of Computer Science</b> <i>Graduate Research Assistant</i>   Advisors: <a href="#">Ziwei Zhu</a> , <a href="#">Antonios Anastasopoulos</a> <u>Project</u> : Exploring Stereotypical Associations in Large Vision-Language Models through generative tasks - image generation and question answering.	<b>Fairfax, USA</b>
<b>Aug 2023</b> <b>May 2023</b>	<b>George Mason University   School of Computer Science</b> <i>Graduate Research Assistant</i>   Advisor: <a href="#">Ziwei Zhu</a> <u>Project</u> : Fairness and Interpretability of Transformer Models.	<b>Fairfax, USA</b>
<b>Jul 2022</b> <b>Jun 2021</b>	<b>University of Technology Sydney   School of Computer Science</b> <i>Visiting Scholar</i>   Advisor: <a href="#">Mukesh Prasad</a> <u>Projects</u> : Disaster Fundraising on Social Media Analysis with the Australian Red Cross, AI integration in aboriginal and indigenous communities' lifestyle, Web information pollution detection (cyberbullying, hate, and offensive speech), Machine learning-based decision-making in clinical vertigo diagnosis.	<b>Sydney, Australia</b>
<b>Aug 2021</b> <b>May 2021</b>	<b>Indian Institute of Management Raipur (IIM-R)</b> <i>Researcher</i>   Advisor: <a href="#">Manojit Chattopadhyay</a> <u>Projects</u> : Cryptocurrencies' price prediction, Analysing causal relationships between stock market behavior and societal events.	<b>Remote</b>
<b>Aug 2022</b> <b>Aug 2019</b>	<b>Delhi Technological University (DTU)</b> <i>Graduate Student Researcher</i>   Advisor: <a href="#">Priyanka Meel</a> <u>Projects</u> : Developing neural network frameworks to detect antisocial web content (fake news, infodemic, misinformation, disinformation, rumor, death hoax, and clickbait), Infodemic impact analysis.	<b>Delhi, India</b>

### Doctoral Publications

- [C.8] **BiasDora: Exploring Hidden Biased Associations in Vision-Language Models** [PDF | Code]  
[Chahat Raj](#), Anjishnu Mukherjee, Aylin Caliskan, Antonios Anastasopoulos, Ziwei Zhu  
Conference on Empirical Methods in Natural Language Processing [EMNLP Findings'24]
- [C.7] **Breaking Bias, Building Bridges: Evaluation and Mitigation of Social Biases in LLMs via Contact Hypothesis** [PDF | Code]  
[Chahat Raj](#), Anjishnu Mukherjee, Aylin Caliskan, Antonios Anastasopoulos, Ziwei Zhu  
AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society [AIIES'24]
- [C.6] **SALSA: Salience-Based Switching Attack for Adversarial Perturbations in Fake News Detection Models** [PDF | Code]  
[Chahat Raj](#)\*, Anjishnu Mukherjee\*, Hemant Purohit, Antonios Anastasopoulos, Ziwei Zhu  
European Conference on Information Retrieval [ECIR'24]

- [C.5] **Global Voices, Local Biases: Socio-cultural Prejudices across Languages** [PDF | Code]  
Anjishnu Mukherjee\*, [Chahat Raj](#)\*, Ziwei Zhu, Antonios Anastasopoulos  
*Conference on Empirical Methods in Natural Language Processing* [EMNLP'23]
- [C.4] **True and Fair: Robust and Unbiased Fake News Detection via Interpretable Machine Learning** [PDF | Code]  
[Chahat Raj](#), Anjishnu Mukherjee, Ziwei Zhu  
*AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society* [AIES'23]

## Pre-Doc Publications

- [J.7] **Machine Learning Models help Differentiate between Causes of Recurrent Spontaneous Vertigo** [PDF]  
Chao Wang, Allison S Young, [Chahat Raj](#), Andrew P Bradshaw, Benjamin Nham, Sally M Rosengren, Zeljka Calic, David Burke, G Michael Halmagyi, Gnana K Bharathy, Mukesh Prasad, Miriam S Welgampola  
*Journal of Neurology* [JoN'24]
- [J.6] **Analyzing Sentiments Towards E-levy Policy Implementation in Ghana using Twitter Data** [PDF]  
Peter Appiahene, Stephen Afrifa, Emmanuel Kyei Akwah, Arjun Choudhry, Inder Khatri, [Chahat Raj](#), Mukesh Prasad  
*International Journal of Information Technology* [IJIT'24]
- [C.3] **Understanding Social Media Engagement in Response to Disaster Fundraising Attempts During Australian Bushfires** [PDF]  
Yohei Nii, [Chahat Raj](#), Muhammad Salman Tiwana, Mahendra Samarawickrama, Simeon Simoff, Tony Jan, Mukesh Prasad  
*International Conference on Intelligent Vision and Computing* [ICIVC'22]
- [C.2] **The Effectiveness of Social Media Engagement Strategy on Disaster Fundraising** [PDF]  
Vivek Velivela, [Chahat Raj](#), Muhammad Salman Tiwana, Raj Prasanna, Mahendra Samarawickrama, Mukesh Prasad  
*International Conference on Information Systems for Crisis Response and Management* [ISCRAM'22]
- [J.5] **People Lie, Actions Don't! Modeling Infodemic Proliferation Predictors among Social Media Users** [PDF]  
[Chahat Raj](#), Priyanka Meel  
*Technology in Society* [TiS'22]
- [S.1] **Is Dynamic Rumor Detection on Social Media Viable? An Unsupervised Perspective** [PDF]  
[Chahat Raj](#), Priyanka Meel  
*In Submission* [2022]
- [J.4] **Cyberbullying Detection: Hybrid Models based on Machine Learning and Natural Language Processing Techniques** [PDF]  
[Chahat Raj](#), Ayush Agarwal, Gnana Bharathy, Bhuva Narayan, Mukesh Prasad  
*Electronics* [Electronics'21]
- [J.3] **Machine Learning Techniques for Differential Diagnosis of Vertigo and Dizziness: A Review** [PDF]  
Varad Kabade, Ritika Hooda, [Chahat Raj](#), Zainab Awan, Allison Young, Miriam Welgampola, Mukesh Prasad  
*Sensors* [Sensors'21]
- [J.2] **ARCNN Framework for Multimodal Infodemic Detection** [PDF]  
[Chahat Raj](#), Priyanka Meel  
*Neural Networks* [NeuNet'21]
- [J.1] **ConvNet Frameworks for Multi-modal Fake News Detection** [PDF]  
[Chahat Raj](#), Priyanka Meel  
*Applied Intelligence* [APIN'21]
- [C.1] **Microblogs Deception Detection using BERT and Multiscale CNNs** [PDF]  
[Chahat Raj](#), Priyanka Meel  
*Global Conference for Advancement in Technology* [GCAT'21]
- [W.1] **An Ensemble-based Multimodal Approach for Coronavirus and 5G Conspiracy Tweet Detection** [PDF]  
[Chahat Raj](#), Mihir P Mehta  
*MediaEval* [MediaEval'20]

## Talks

### “A Psychological View to Social Biases in LLMs”

› SouthNLP @ Emory University

April 2024 (Georgia, USA)

### “Cyberbullying Monitoring”

› ML4AU Research Showcase Event

2021 (Remote)

## Awards

---

**CAHMP Fellowship @ George Mason University, 2024** For summer research on social bias assessment in visual question answering.

**Best Paper Award @ MASC-SLL 2024, Johns Hopkins University** For the paper ‘A Psychological View to Social Bias in LLMs: Evaluation and Mitigation’.

**Summer Graduate Research Award @ George Mason University, 2023** For performing research on ‘Explainability and Robustness of Debiasing Approaches for Transformers’.

**Travel Grant by AAAI/ACM AIES 2023** Funded \$1000 to travel to Montreal, Canada to attend the AIES conference.

**Research Excellence Award @ Delhi Technological University 2023 (×2) & 2022** Received three Commendable Research Awards with a total cash prize of INR 150k by DTU, Delhi, for the research published in reputed scientific journals.

**Graduate Scholarship @ Delhi Technological University 2019** Received the AICTE scholarship of INR 297k for qualifying the GATE exam

## Teaching

---

**CS 108 Introduction to Computer Programming, GMU** *Teaching Assistant* Spring’24

› Responsible for teaching labs, one-to-one student tutoring, creating lab modules, grading, and exam invigilation.

**CS 478 Natural Language Processing, GMU** *Teaching Assistant* Fall’23

› Delivered ad-hoc lectures, student tutoring, teaching labs, and assignment grading.

**COMP 502 Mathematical Foundations of Computing, GMU** *Teaching Assistant* Fall’23

› Responsible for one-to-one student tutoring sessions, and assignment grading.

**CS 112 Introduction to Python Programming, GMU** *Teaching Assistant* Fall’22, Spring’23

› Responsible for teaching labs, one-to-one student tutoring, creating lab modules, assignment grading, and exam invigilation.

## Academic Service

---

**Reviewer** ACM Transactions on Intelligent Systems and Technology 2024, LTEDI @ EACL 2024, TrustNLP 2024, CSCW 2024, NeurIPS 2023, Elsevier & Springer Journal papers

**Sub-Reviewer** EMNLP’24, AIES’24

## Student Mentorship

---

› Mamnuya Rinki [MS, GMU]

› Diwita Banerjee [MS, GMU]

› Samyak Jain [MS, GMU]