Gaurav Verma

Computer Science Ph.D. Student
Georgia Institute of Technology
Atlanta, GA 30308

Email: gverma9@gatech.edu
Webpage: gaurav22verma.github.io/

Education

- Ph.D. in Computer Science, **Georgia Institute of Technology**   (Jan 2021 –)
  Research Interests: Robust Multimodal Learning, NLP, Comp. Social Science
  Advisor: Prof. Srijan Kumar; *GPA*: 4.0/4.0
  Minor: Science and Technology Policy (School of Public Policy @ Georgia Tech)

- Bachelor of Technology, **Indian Institute of Technology Kanpur**   (July 2014 – May 2018)
  Major in Electrical Engineering and Minor in Machine Learning (CSE)
  Distinction and Academic Excellence Awards

Experience

- Graduate Research Assistant at **Georgia Institute of Technology**   (Jan 2021– present)
  Topics: Robust Multimodal Machine Learning, Natural Language Generation

- Research Intern at **Microsoft Research**, Canada   (May 2023 - Aug 2023)
  Mentors: Drs. Alexandra Olteanu, Su Lin Blodgett, & Koustuv Saha
  Topic: Understanding User Conceptualization of and Behavior with LLM-based Search

- Research Scientist Intern at **Adobe Research**, Maryland, USA   (May 2022 - Aug 2022)
  Mentors: Drs. Ani Nenkova & Ryan A. Rossi
  Topic: Modeling the Visualness of Text using Large Vision-Language Models

- Research Fellow at **Adobe Research**, India   (June 2018 – Nov 2020)
  Topics: Multimodal Learning, Natural Language Processing, and Information Retrieval

- Research Intern at **Adobe Research**, India   (May 2017 – July 2017)
  Mentor: Dr. Vishwa Vinay; *Topic*: Natural Language Processing for Semantic Search

Awards & Honors

- J.P. Morgan Chase AI Research Ph.D. Fellowship   2023
  - "To recognize and enable the next generation of leading AI researchers"

- Snap Research Fellowship   2022
  - "To support most driven graduate CS students across the world"
  – Award amount: $10,000 | One of 12 recipients out of over 150 applicants

- Adobe Research Ph.D. Fellowship Finalist   2022
  - "Recognizes outstanding students globally carrying out exceptional research"
  - One of the 30 candidates shortlisted for the final round

- College of Computing Rising Star Doctoral Student Research Award (@ Georgia Tech)   2022
  - "For meritorious research contributions during the first semester/year"
  - The annual award recognizes one Ph.D. student among all first-year CoC doctoral students

- Best Paper Runner-up Award at WISE 2021   2021

- AAAI ICWSM-2021 Best Reviewer Award 2021

- Awarded Outstanding Grade (Top ∼1%) for Undergraduate Thesis at IIT Kanpur 2018

- Academic Excellence Awards by IIT Kanpur 2015 & 2016

- Bronze Medalist at the 6th International Earth Science Olympiad 2012
  - Student Leader of the Indian Team (4 students and 2 professors) at Olavarría, Argentina
  - Ranked first in the national round after a three-week training camp at Goa University
  - "Indian students out to prove their mettle at IESO in Argentina" - The Times of India

- Kishore Vaigyanik Protsahan Yojana (KVPY) Scholar 2012
  - Awarded by Indian Institute of Science (IISc) to students with an aptitude for research
  - Acceptance rate of $< 0.5\%$ out of $\sim 100,000$ applicants

- Infosys Award by HBCSE and Tata Institute of Fundamental Research, Mumbai 2012

- Awarded National Talent Search Scholarship 2010
  - Conducted by NCERT to identify and nurture talented students
  - Acceptance rate of $< 0.1\%$ out of $\sim 1,000,000$ applicants

- Commendation by Air Force School, Pathankot for voluntary teaching services
  - Taught science and mathematics to students of classes 9th through 12th (equivalent to high school students) during summer and winter vacations

Publications

Journals

2. *Examining the Impact of Sharing COVID-19 Misinformation Online on Mental Health*
   Gaurav Verma, Ankur Bhardwaj, Talayeh Aledavood, Munmun De Choudhury, Srijan Kumar
   In Scientific Reports. 2022. (*IF: 4.9*) (https://www.nature.com/articles/s41598-022-11488-y)
   - In Press: "Why Tackling Medical Misinformation Is Key To Pandemic Recovery" - **Forbes** [link]
   - In "Misinformation: Yep, still a problem' - by Prof. Timothy Caulfield in **Healthy Debate** [link]
   - Invited presentation at the Oxford Internet Policy and Politics Conference [link]

1. *Large-Scale Text Analysis Using Generative Language Models: Discovering Public Values in AI Patents*
   Sergio Pelaez*, **Gaurav Verma***, Barbara Ribeiro, Philip Shapira (* = equal contribution)
   (To Appear) *In* Quantitative Science Studies (QSS). 2024. (*IF: 5.1*) [preprint] [code]

Peer-reviewed Conferences

16. *Learning the Visualness of Text Using Large Vision-Language Models* [arXiv] [webpage]
    Gaurav Verma, Ryan A. Rossi, Chris Tensmeyer, Jiuxiang Gu, Ani Nenkova
    The 2023 Conf. on Empirical Methods in Natural Language Processing (**EMNLP 2023**).
    - Also presented at *Creative AI Across Modailities* Workshop @ AAAI 2023

15. *Cross-Modal Attribute Insertions for Assessing the Robustness of Vision-and-Language Learning* [arXiv] [code]
    Shivaen Ramshetty*, **Gaurav Verma***, Srijan Kumar (* = equal contribution)
    The 61st Annual Meeting of the Association for Computational Linguistics (**ACL 2023**).

14. *Robustness of Prompt-Based Few-Shot Learning for Natural Language Understanding* [arXiv] [code]
    Sarath Nookala*, **Gaurav Verma***, Subhabrata Mukherjee, Srijan Kumar (* = equal contribution)
    The 61st Annual Meeting of the Association for Computational Linguistics (**ACL 2023 Findings**).

13. *Robustness of Fusion-based Multimodal Classifiers to Cross-Modal Content Dilutions* [arXiv] [code] [web]
Gaurav Verma, Vishwa Vinay, Ryan A. Rossi, Srijan Kumar
The 2022 Conf. on Empirical Methods in Natural Language Processing (**EMNLP 2022**).
(Oral presentation; $< 4\%$ submitted papers)

12. *Minority Stress Experienced by LGBTQ Online Communities During the COVID-19 Pandemic* [arXiv]
Yunhao Yuan, **Gaurav Verma**, Barbara Keller, Talayeh Aledavood
AAAI International Conference on Web and Social Media (**ICWSM 2023**). (Acceptance rate: $\sim$20%)

11. *Overcoming Language Disparity in Online Content Classification with Multimodal Learning* [pdf] [webpage]
Gaurav Verma, Rohit Mujumdar, Zijie J. Wang, Munmun De Choudhury, Srijan Kumar
AAAI International Conference on Web and Social Media (**ICWSM 2022**). (Acceptance rate: 22.0%)
- Covered in 'BERT for Social Sciences & Humanities' Tutorial @ ICWSM 2022 [webpage, slides]
- Mentioned in Prof. Alice Oh's keynote on 'The importance of multiple languages and multiple cultures in NLP research' @ ICWSM 2022

10. *Characterizing, Predicting, and Detecting Online Ban Evasion* [pdf] [dataset] [video]
Manoj Niverthi*, **Gaurav Verma***, Srijan Kumar (* = equal contribution)
ACM Web Conference 2022 (**TheWebConf 2022**). (Acceptance rate: 17.7%)
- In deployment on **Wikipedia** to support moderators in real-time detection of ban evasion
- Invited research summary at the Montreal AI Ethics Institute [link]
- Invited presentation in Web Search and Text Mining course (CSE 6240) @ Georgia Tech
- "Computer Scientists Build, Test, and Present Model to Curb Online Ban Evasion" [GT CoC Article]

9. *Non-Linear Consumption of Videos Using a Sequence of Personalized Multimodal Fragments* [pdf]
Gaurav Verma, Trikay Nalamada, Keerti Harpavat, Pranav Goel, Aman Mishra, Balaji Vasan Srinivasan
The 26th International Conference on Intelligent User Interfaces (**IUI 2021**). (Acceptance rate: 25%)

8. *DRAG: Director-Generator Language Modelling Framework for Non-Parallel Author Stylized Rewriting* [arXiv]
Hrituraj Singh, **Gaurav Verma**, Aparna Garimella, Balaji Vasan Srinivasan
The 16th Conference of the European Chapter of the Association for Computational Linguistics (**EACL 2021**). (Acceptance rate: 24.7%)

7. *Detecting Document Versions and Their Ordering In a Collection* [pdf]
Natwar Modani, Anurag Maurya, **Gaurav Verma**, Inderjeet Nair, Vaidehi Patil, Anirudh Kanfade
The 22nd International Conference on Web Information Systems and Engineering (**WISE 2021**). (Acceptance rate: 24%) [⋆ Best Paper Runner-up Award]

6. *Incorporating Stylistic Lexical Preferences in Generative Language Models* [arXiv]
Hrituraj Singh, **Gaurav Verma**, Balaji Vasan Srinivasan
The 2020 Conference on Empirical Methods in Natural Language Processing (**EMNLP Findings 2020**).
(Acceptance rate: 38.0 %)

5. *Using Image Captions and Multitask Learning for Recommending Query Reformulations* [arXiv]
Gaurav Verma, Vishwa Vinay, Sahil Bansal, Shashank Oberoi, Makkunda Sharma, Prakhar Gupta
The 42nd European Conference on Information Retrieval (**ECIR 2020**). (Acceptance rate: 26.6 %)

4. *Generating Need-Adapted Multimodal Fragments* [pdf]
Gaurav Verma, Suryateja BV, Samagra Sharma, Balaji Vasan Srinivasan
The 25th International Conference on Intelligent User Interfaces (**IUI 2020**). (Acceptance rate: 23.4 %)
- "Innovative ideas that evolve the way brands interact with their customers" – **Adobe Summit 2020**
- "Adobe tests an AI recommendation tool for headlines and images" – **TechCrunch**
- Invited presentation in 'Out of India' track of India HCI 2020

3. *Adapting Language Models for Non-parallel Author-Stylized Rewriting* [arXiv]
Bakhtiyar Syed, **Gaurav Verma**, Balaji Vasan Srinivasan, Anandhavelu Natarajan, Vasudeva Varma
The 34th AAAI Conference on Artificial Intelligence (**AAAI 2020**). (Acceptance rate: 20.6 %)

2. *Learning Affective Correspondence between Music and Image* [arXiv]
   Gaurav Verma, Eeshan Gunesh Dhekane, Tanaya Guha
   The 44th International Conference on Acoustics, Speech and Signal Processing (**ICASSP 2019**). (Acceptance rate: 46.5 %)

1. *Stuck? No Worries!: Task-aware Command Recommendation and Proactive Help for Analysts* [arXiv]
   Aadhavan M. Nambhi*, Bhanu Prakash Reddy*, Aarsh Prakash Agarwal*, **Gaurav Verma**, Harvineet Singh, Iftikhar Ahamath Burhanuddin (* = equal contribution)
   The 27th Conference on User Modelling, Adaptation And Personalization (**UMAP 2019**). (Acceptance Rate: 20.3 %)

Workshops, Posters, & Data Challenges (*Lightly Peer-reviewed & Non-Archival*)

5. *How does artificial intelligence contribute to public values? A large-scale patent analysis.*
   Sergio Pelaez, **Gaurav Verma**, Barbara Ribeiro, Philip Shapira
   - 4POINT0 (Partnership for the Organisation of Innovation and New Technologies) International Consortium @ Polytechnique Montréal (2022) — (Presentation by Sergio Pelaez)
   - Accepted talk at the 9th Atlanta Conference on Science and Innovation Policy (by Sergio Pelaez) (2023)
   - Invited talk at the 13th Global TechMining Conference (by Sergio Pelaez) (2023)

4. *The Paradox of Integration: Interdisciplinary Research and Disciplinary Evaluation*
   Robert Ward, **Gaurav Verma**, Seokkyun Woo
   6th Summer School on Data and Algorithms for Science, Technology & Innovation Studies (2021) (Presentation by Robert Ward)

3. *Modeling Causal Impact of Textual Style on a Targeted Goal* [pdf]
   Gaurav Verma, Balaji Vasan Srinivasan, Shiv Kumar Saini, Niyati Chhaya
   In WWW'20 Companion: Companion Proceedings of The Web Conference 2020 (**WWW 2020**).

2. *Identification and Analysis of Abusive Text Using Ensemble of Classifiers* [arXiv]
   Gaurav Verma, Niyati Chhaya, Vishwa Vinay
   Safety Data Challenge at The 14th International AAAI Conference on Weblogs and Social Media. ICWSM Data Challenge 2020.

1. *Semi-Supervised Learning for Identification of Informative COVID-19 English Tweets* [arXiv]
   Abhilasha Sancheti*, Kushal Chawla*, **Gaurav Verma** (* = equal contribution)
   Shared Task 2, Workshop on Noisy User-Generated Text at EMNLP 2020.

Patents (*Granted & Applications*)

10. *Identifying Visual Text Using Vision-Language Models* — Jiuxiang Gu, Ryan Rossi, **Gaurav Verma**, Christopher Tensmeyer, Ani Nenkova — US Patent Application (In filing) | Adobe Inc.

9. *Personalizing Videos with Non-Linear Playback* [pdf] — **Gaurav Verma**, Trikay Nalamada, Keerti Harpavat, Pranav Goel, Aman Mishra, Balaji Vasan Srinivasan — US Patent No. 11,670,085 | Adobe Inc.

8. *Content Fragments Aligned to Content Criteria* [pdf] — **Gaurav Verma**, Suryateja BV, Samagra Sharma, Balaji Vasan Srinivasan — US Patent No. 11,308,146 | Adobe Inc.

7. *Automatically Associating Context-based Sounds with Text* [pdf] — **Gaurav Verma**, Vishwa Vinay, Sneha Chowdary, Siddharth Sahay, Mitansh Jain — US Patent No. 11,727,913 | Adobe Inc.

6. *Goal-driven Authoring Assistance via Causal Stylistic Prescriptions* [pdf] — **Gaurav Verma**, Balaji Vasan Srinivasan, Shiv Kumar Saini, Niyati Chhaya — US Patent No. 11,062,085 | Adobe Inc.

5. *Task-aware Command Recommendations and Proactive Help* [pdf] — **Gaurav Verma**, IA Burhanuddin, Harvineet Singh, BP Reddy, AM Nambhi, AP Agarwal — US Patent Application 16/454,683 | Adobe Inc.

4. *Systems for Generating Indications of Relationships Between Documents* [pdf] — Natwar Modani, Vaidehi Patil, Inderjeet Nair, **Gaurav Verma**, A Maurya, A Kanfade — US Patent App. 17/534,744 | Adobe Inc.

3. *Stylistic Text Rewriting for a Target Author* [pdf] — Balaji Vasan Srinivasan, **Gaurav Verma**, Bakhtiyar Hussain Syed, Anandhavelu Natarajan — US Patent No. 11,157,693 | Adobe Inc.

2. *Digital Document Update Using Static and Transient Tags* [pdf] — Vishwa Vinay, Sopan Khosla, Sanket V Mehta, Sahith Thallapally, **Gaurav Verma** — US Patent No. 10,846,466 | Adobe Inc.

1. *Digital Document Update* [pdf] — Vishwa Vinay, Sopan Khosla, Sanket V Mehta, Sahith Thallapally, **Gaurav Verma** — US Patent No. 10,489,498 | Adobe Inc

Teaching, Academic Service, & Mentoring

Teaching

- **Head Teaching Assistant** for CSE 6240 @ Georgia Tech   (Spring 2023)
  - Course details: *Web, Search, and Text Mining* with Prof. Srijan Kumar
  - Led a team of teaching assistants for a graduate-level course with about 100 students, created and graded homework and exams, provided feedback to the students, and conducted regular office hours
  - Delivered in-person lectures on 'Large Language Models' and 'Graph Neural Networks'

Academic Service

- **PC Member/Reviewer**
  - *Journals*: ACM Transactions on Computer-Human Interaction (**TOCHI**), IEEE Transactions on Pattern Analysis and Machine Intelligence (**PAMI**), Data Mining & Knowledge Discovery
  - *Conferences*: **NeurIPS** (2023: Ethics reviewer), **AAAI** (2022, 2023, 2024), **EMNLP** (2021, 2022, 2023), **ACL** (2021, 2023), **TheWebConf (WWW)** (2022, 2023, 2024), **KDD** (2023) **ICLR** (2022), **CHI** (2023: ⋆ Special Recognition for Outstanding Reviews), **ICWSM** (2021: ⋆ Best Reviewer Award, 2022, 2023, 2024), **ACII** (2021)
  - *Workshops*: **Creative AI Across Modalities @ AAAI** (2023), **Wiki Workshop @ WWW** (2022)

- **Student Volunteer** for **NAACL 2021**

Mentoring & Misc.

- **Mentoring experience with student collaborators**
  - Mingxiao (Michelle) Song, BS CS @ Georgia Tech (September 2023 –)
  - Ishwarya Sivakumar, MS CS @ Georgia Tech (September 2023 –)
  - Rynaa Grover, MS CS @ Georgia Tech (November 2022 –)
  - Shivaen Ramshetty, MS CS @ Georgia Tech → Toyon Research Corp. (August 2021 – Jan 2023)
        Topic: robust multimodal learning; resulted in a co-authored ACL 2023 paper
  - Sarath Nookala, MS CS @ Georgia Tech → Meta (Mar 2022 – Jan 2023)
        Topic: robust few-shot NLP; resulted in a co-authored ACL 2023 (Findings) paper
  - Manoj Niverthi, BS CS → MS CS @ Georgia Tech (Jan 2021 – April 2023)
        Topics: machine learning, ban evasion; resulted in a co-authored WebConf 2022 paper
  - Jio Oh, Exchange student @ Georgia Tech (BS CS @ KAIST) (Feb 2022 – May 2022)
  - Zhen Jiang, BS CS @ Georgia Tech → UC Berkeley (July 2021 – May 2022)
  - Ankur Bhardwaj, MS Data Science @ Georgia Tech → Walmart Labs, US (June 2020 – April 2021)
        Topics: causal inference, well-being; resulted in a co-authored Scientific Reports paper
  - Co-mentored 14 UG students in Adobe Research India's Summer Internship program (2018-2020)

- Reviewer for President's Undergraduate Research Awards at Georgia Tech   (2023)

- Volunteer for Military@Microsoft and Asians@Microsoft Employee Resource Groups   (May-Aug 2023)

- Mentor for Georgia Tech's Graduate Application Support (GT-GAS) program to support underrepresented applicants in the field of computer science (2022)

- Machine Learning Tutor for Adobe's Employee Program (2018-2019)

- Member of Recruitment Team for Interns at Adobe Research, India (2018-2020)

- Voluntary teaching at Air Force School, Pathankot and Prayas, IIT Kanpur (2016-2017)