MAHAMMED KAMRUZZAMAN

Portfolio Google Scholar
kamruzzaman1@usf.edu LinkedIn GitHub

EDUCATION

Ph.D. in Computer Science and Engineering *Aug 2022-Present*
University of South Florida, Tampa
Advisor: Gene Louis Kim

M.Sc. in Computer Science (concurrent) *Aug 2022-Dec 2024*
University of South Florida, Tampa

B.Sc. in Electrical and Computer Engineering *Feb 2016- Jan 2021*
Rajshahi University of Engineering & Technology, Bangladesh

EXPERIENCE

Graduate Assistant Aug 2022–Present
University of South Florida, Tampa

- Conduct research on fairness and bias in large language and vision–language models, focusing on social, cultural, brand, and disability-related disparities in AI predictions and reasoning.
- Design and evaluate prompt-based and model-agnostic frameworks to quantify bias and improve equity in NLP and multimodal tasks such as emotion recognition and harmful language detection.
- Assisted in teaching graduate courses in Human-Computer Interaction and Operating Systems, mentoring students on data-driven user research and usability-focused design.

AI Engineering Intern Aug 2026–Present
Metadata Inc., San Francisco, CA

Mentors: Gil Allouche and Manikandan Sekar

- Design and evaluate LLM-based agent workflows that combine API/tool calling, structured outputs, evidence validation, and guardrails to automate research and data-processing tasks reliably.
- Evaluate and incorporate data sources to enrich Metadata's identity graph, assessing identity resolution, person-to-company relationships, and identifier quality, while proposing new data signals and sources to improve entity coverage and matching accuracy.

PUBLICATIONS

- **M. Kamruzzaman**, S. Das, and G. L. Kim, "Demographic Prompting at Scale: When More Attributes Hurt LLM–Human Agreement," In Proceedings of the 2026 Conference on Empirical Methods in Natural Language Processing. ***(EMNLP 2026 Main– Accepted)***
- **M. Kamruzzaman**, H. M. Nguyen, N. Hassan, and G. L. Kim, "A Woman is More Culturally Knowledgeable than A Man?: The Effect of Personas on Cultural Norm Interpretation in LLMs," in Proceedings of the First Workshop on Multilingual Multicultural Evaluation, 2026. ***(MME Workshop, EACL 2026)***
- K. E. Miandoab*, **M. Kamruzzaman***, A. Gharooni, G. L. Kim, V. Sarathy, and N. Mehrabi, "Breaking the Benchmark: Revealing LLM Bias via Minimal Contextual Augmentation," in Proceedings of the Fifteenth Language Resources and Evaluation Conference. ***(LREC 2026 Main)*** *(* Equal contribution)*

- **M. Kamruzzaman**, A. A. Monsur, G. L. Kim, and A. Chhabra, "From Anger to Joy: How Nationality Personas Shape Emotion Attribution in Large Language Models," In Proceedings of the International Joint Conference on Natural Language Processing & Asia-Pacific Chapter of the Association for Computational Linguistics. ***(AACL 2025 Main)***
- **M. Kamruzzaman** and G. L. Kim, "Exploring Changes in Nation Perception with Nationality-Assigned Personas in LLMs," In Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing, pages 3661–3679. ***(EMNLP 2025 Main)***
- **M. Kamruzzaman**, A. C. Curry, A. Curry, and F. Plaza, "Seeing Race, Feeling Bias: Emotion Stereotyping in Multimodal Language Models," in Findings of the Association for Computational Linguistics: EMNLP 2025, pages 7317–7351. ***(EMNLP 2025 Findings)***
- **M. Kamruzzaman** and G. L. Kim, "The Impact of Name Age Perception on Job Recommendations in LLMs," in Findings of the Association for Computational Linguistics: ACL 2025, pages 15033–15058, Vienna, Austria. ***(ACL 2025 Findings)***
- **M. Kamruzzaman**, A. A. Monsur, S. Das, E. Hassan, and G. L. Kim, "BanStereoSet: A Dataset to Measure Stereotypical Social Biases in LLMs for Bangla," in Findings of the Association for Computational Linguistics: ACL 2025, pages 3450–3460, Vienna, Austria. ***(ACL 2025 Findings)***
- **M. Kamruzzaman** and G. L. Kim, "Evaluating the Impact of Racial Cues on MLLMs Judgements of Politeness and Offensiveness," IEEE/CVF International Conference on Computer Vision (ICCV) Workshops. ***(FAILED Workshop, ICCV 2025)***
- **M. Kamruzzaman** and G. L. Kim, "Prompting Techniques for Reducing Social Bias in LLMs through System 1 and System 2 Cognitive Processes," In Proceedings of the 15th International Conference on Recent Advances in Natural Language Processing, pages 511–520, Varna, Bulgaria. ***(RANLP 2025 Main)***
- **M. Kamruzzaman** and G. L. Kim, "The Impact of Disability Disclosure on Fairness and Bias in LLM-Driven Candidate Selection," The International FLAIRS Conference Proceedings, 38(1). ***(FLAIRS 2025)***
- **M. Kamruzzaman**, "Investigating and Mitigating Undesirable Biases in Large Language Models," In Proceedings of the AAAI Conference on Artificial Intelligence, volume 39, pages 29273–29274. ***(Doctoral Consortium, AAAI 2025)***
- **M. Kamruzzaman**, H. M. Nguyen, and G. L. Kim, "Global is Good, Local is Bad?: Understanding Brand Bias in LLMs," In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, pages 12695–12702. ***(EMNLP 2024 Main)***
- **M. Kamruzzaman**, M. M. I. Shovon, and G. L. Kim, "Investigating Subtler Biases in LLMs: Ageism, Beauty, Institutional, and Nationality Bias in Generative Models," in Findings of the Association for Computational Linguistics: ACL 2024, pages 8940–8965, Bangkok, Thailand. ***(ACL 2024 Findings)***
- **M. Kamruzzaman** and G. L. Kim, "Efficient Sentiment Analysis: A Resource-Aware Evaluation of Feature Extraction Techniques, Ensembling, and Deep Learning Models," in Proceedings of the 11th International Workshop on Natural Language Processing for Social Media, Association for Computational Linguistics, Nov. 2023, pp. 9–20. ***(SocialNLP Workshop, AACL 2023)***
- **M. Kamruzzaman**, M. M. I. Shovon, and G. Kim, "BanMANI: A Dataset to Identify Manipulated Social Media News in Bangla," in Proceedings of the Workshop on Computational Terminology in NLP and Translation Studies (ConTeNTS), Incorporating the 16th Workshop on Building and Using Comparable Corpora (BUCC), pp. 51–58. ***(BUCC Workshop, RANLP 2023)***

- **M. Kamruzzaman**, M. Hossain, M. R. I. Imran, and S. C. Bakchy, "A Comparative Analysis of Sentiment Classification Based on Deep and Traditional Ensemble Machine Learning Models," in 2021 International Conference on Science & Contemporary Technologies (ICSCT). IEEE, 2021, pp. 1–5. ***(ICSCT 2021)***
- **M. Kamruzzaman**, S. C. Bakchy, T. Hasan, M. M. I. Shovon, and M. Hossain, "An Intense Investigation on Deep and Traditional Machine Learning Approaches for Better Sentiment Classification," in 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI). IEEE, 2021, pp. 1–6. ***(ACMI 2021)***
- T. Hasan, A. Matin, **M. Kamruzzaman**, S. Islam, and M. O. F. Goni, "A Comparative Analysis of Feature Extraction Methods for Human Opinion Grouping Using Several Machine Learning Techniques," in 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE). IEEE, 2020, pp. 272–275. ***(WIECON-ECE 2020)***
- M. M. I. Shovon, **M. Kamruzzaman**, and M. Kundu, "Recognition of Handwritten Bangla Number Using Multi-Layer Convolutional Neural Network," in 2020 IEEE Region 10 Symposium (TENSYMP). IEEE, 2020, pp. 783–786. ***(TENSYMP 2020)***
- H. M. Nguyen, **M. Kamruzzaman**, A. Chhabra, and G. L. Kim, "Token-Level Diagnosis of Sycophancy in LLMs with Attribution-Guided Steering," 2026. ***[Under Review]***

AWARD AND SCHOLARSHIP

Chih Foundation Research and Publication Award — *University of South Florida*
(Top 3 awardees university-wide)

Spirit of Innovation Award — *University of South Florida*

Travel Scholarship — *AAAI-2025*

PRESS COVERAGE

Your most important customer may be AI — Feb, 2025
MIT Technology Review

Interview with Mahammed Kamruzzaman — June, 2025
AIhub

CONFERENCE REVIEWING

ACL Rolling Review (ARR) — February 2025 (ACL 2025)

ACL Rolling Review (ARR) — May 2025 (EMNLP 2025)

6th Workshop on Gender Bias in Natural Language Processing — co-located with ACL-2025

8th AAAI/ACM Conference on AI, Ethics, and Society — 2025

STUDENT ADVISING AND MENTORING

- **Hieu Nguyen**, Master's Student, University of South Florida
- **Jason Rudinsky**, Master's Student, University of South Florida
- **Shrabon Kumar Das**, Undergraduate Student, University of South Florida
- **Anubhav Gupta**, Undergraduate Student, University of South Florida *(mentored through the UR2PhD Graduate Student Mentor Training Course)*

- **Otavio Rosa**, Undergraduate Student, University of South Florida *(mentored through the UR2PhD Graduate Student Mentor Training Course)*

TEACHING ASSISTANT

- **COP4530: Data Structures** *Summer 2026*
  Undergraduate Course, University of South Florida
- **COP6611: Operating Systems** *Spring 2023, Fall 2023, Spring 2024*
  Graduate Course, University of South Florida
- **COP2512: IT Programming Fundamentals** *Fall 2022*
  Undergraduate Course, University of South Florida
- **CIS6930: Human Computer Interaction** *Fall 2022*
  Graduate Course, University of South Florida

GUEST LECTURES

- **CAI 4002: Introduction to Artificial Intelligence** *Fall 2025*
  Guest Lecture, University of South Florida