

Kshitish Ghate

kshitishghate.github.io
kghate@cs.washington.edu

Paul G. Allen School of Computer Science and Engineering
University of Washington
185 E Stevens Way NE, Seattle, WA 98195

◇ EDUCATION

- 2025 – Present **University of Washington**, Seattle, WA
PH.D. IN COMPUTER SCIENCE & ENGINEERING
Advisors: Aylin Caliskan, Tadayoshi Kohno
- 2023 – 2025 **Carnegie Mellon University**, Pittsburgh, PA
MASTERS IN LANGUAGE TECHNOLOGIES
Advisor: Mona Diab
Cumulative GPA: 4.10/4.33
Relevant Coursework: Advanced NLP, Introduction to Machine Learning (graduate level), Multi-modal ML, On-Device ML, LLM Applications, Question and Answering
- 2018 – 2023 **Birla Institute of Technology and Science**, Pilani, Gos, India
BACHELOR OF ENGINEERING IN COMPUTER SCIENCE AND MASTER OF SCIENCE IN ECONOMICS
Cumulative GPA: 8.99/10
Relevant Coursework: Artificial Intelligence, Data Structures and Algorithms, Database Management Systems, Operating Systems, Foundations of Data Science, Game Theory, Machine Learning, Object Oriented Programming, Probability and Statistics

◇ PUBLICATIONS

— Select Conference and Journal Publications —

- 2025 **Kshitish Ghate**, Tessa Charlesworth, Mona T. Diab, and Aylin Caliskan. “Biases Propagate in Encoder-based Vision-Language Models: A Systematic Analysis From Intrinsic Measures to Zero-shot Retrieval Outcomes.” *In Findings of the Association for Computational Linguistics: ACL 2025*, pages 18562–18580, Vienna, Austria. Association for Computational Linguistics.
- 2025 Jingwen Cheng, **Kshitish Ghate**, Wenyue Hua, William Yang Wang, Hong Shen, and Fei Fang. “REALM: A Dataset of Real-World LLM Use Cases.” *In Findings of the Association for Computational Linguistics: ACL 2025*, pages 8331–8341, Vienna, Austria. Association for Computational Linguistics.
- 2025 **Kshitish Ghate***, Isaac Slaughter*, Kyra Wilson, Mona Diab, and Aylin Caliskan. “Intrinsic Bias is Predicted by Pretraining Data and Correlates with Downstream Performance in Vision-Language Encoders.” *In Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 2899–2915, Albuquerque, New Mexico. Association for Computational Linguistics.
- 2024 Tessa Charlesworth, **Kshitish Ghate**, Aylin Caliskan, and Mahzarin R. Banaji. “Extracting intersectional stereotypes from embeddings: Developing and validating the Flexible Intersectional Stereotype Extraction procedure.” *PNAS nexus* 3, no. 3 (2024).

2024 Alfiya M. Shaikh, Hrithik Nambiar*, **Kshitish Ghate***, Swarnali Banik, Sougata Sen, Surjya Ghosh, Vaskar Raychoudhury, Niloy Ganguly and Snehanshu Saha. "Self-SLAM: A Self-supervised Learning Based Annotation Method to Reduce Labeling Overhead." *In Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 123-140). Cham: Springer Nature Switzerland.

— Select Workshop Publications —

2024 **Kshitish Ghate**, Arjun Choudhry, and Vanya Bannihatti Kumar. "Evaluating Gender Bias in Multilingual Multimodal AI Models: Insights from an Indian Context." *In Proceedings of the 5th Workshop on Gender Bias in Natural Language Processing (GeBNLP)*, pp. 338-350. 2024.

2024 Nishant Subramani*, **Kshitish Ghate***, and Mona Diab. "Evaluating Personal Information Parroting in Language Models." *Accepted to the 4th Workshop on Trustworthy Natural Language Processing*.

◇ TEACHING

Fall 2024 LARGE LANGUAGE MODELS: METHODS AND APPLICATIONS (11-667), Teaching Assistant, CMU

Spring 2022 APPLIED ECONOMETRICS, Teaching Assistant, BITS Pilani

Fall 2021 SECURITIES ANALYSIS AND PORTFOLIO MANAGEMENT, Teaching Assistant, BITS Pilani

◇ SKILLS

Languages PYTHON, R, STATA, C/C++, MATLAB, JAVA (INTERMEDIATE), SQL (INTERMEDIATE)

Toolkits/Cloud PYTORCH, TENSORFLOW, KERAS, SCIKIT-LEARN, PANDAS, NUMPY, MATPLOTLIB, GCP, AWS, AZURE, PYSPARK (INTERMEDIATE)

◇ INTERNSHIPS & ASSISTANTSHIPS

08/23 – present **Carnegie Mellon University**, Pittsburgh, PA

GRADUATE RESEARCH ASSISTANT, Language Technologies Institute

Advisors: Mona Diab, Maarten Sap

- Leading a project on pluralistic alignment to develop AI systems that align with diverse, potentially competing values, ensuring tailored, equitable responses for various stakeholders.
- Evaluated novel confidence-aware calibration strategies across four open-source LLMs, achieving state-of-the-art accuracy and calibration on the MediQ interactive medical QA benchmark.
- Developed a risk taxonomy for Personal Information (PI) memorization in large language models and improved existing detectors (WIMBD) with a superior suite, achieving 90% better FPR.

05/24 – 08/24 **University of Washington**, Seattle, WA

RESEARCH SCIENTIST INTERN, Information School

Advisors: Aylin Caliskan, Mona Diab

- Designed a harm reduction framework for reducing bias and hallucinations in LLM responses used for clinical decision-making by balancing invariance and adaptability across demographics.
- Investigated the relationship between intrinsic social biases in 131 CLIP models, their pretraining factors and downstream performance on the VTAB+ benchmark.

- 11/23 – 07/24 **Carnegie Mellon University, Pittsburgh, PA**
 GRADUATE RESEARCH ASSISTANT, Software and Societal Systems Department
 Advisors: Fei Fang
 Developed NewsPanda++, a multisource, multimodal system for detecting environmental conservation events, deployed by WWF India and Nepal for monitoring and decision-making.
- 01/23 – 08/23 **University of Washington, Seattle, WA**
 UNDERGRADUATE THESIS RESEARCH INTERN, Information School
 Advisors: Aylin Caliskan, Tessa Charlesworth
 Demonstrated the propagation of intrinsic social biases to downstream zero-shot text-to-image and image-to-text retrieval settings in vision-language models such as CLIP, BLIP-2.
- 03/22 – 08/23 **Harvard University, Cambridge, MA**
 RESEARCH ASSISTANT, Implicit Social Cognition Lab
 Advisors: Mahzarin Banaji, Tessa Charlesworth, Aylin Caliskan
 Worked on the Flexible Intersectional Stereotype Extraction (FISE) method to quantify intersectional stereotypes in large language corpora, revealing key insights into social group dynamics.
- 08/22 – 12/22 **Amazon, Bangalore, India**
 APPLIED SCIENTIST INTERN, Alexa – AI Natural Language Understanding
 Supervisors: Anurag Dwarakanath, Anjali Shenoy
 Implemented a novel Curriculum Learning training methodology to address problem of classifying long tail data and achieved 5% improvement in Slot F1 and Intent Classification.
- 08/22 – 12/22 **Cambridge Judge Business School, Cambridge, UK**
 RESEACH INTERN, Cambridge Centre for Alternative Finance
 Supervisors: Tania Ziegler,
 Automated the data extraction process and streamlined the workflow associated with data analysis; contributed to the Global COVID-19 Fintech Market and Impact Resilience Study.
- 08/20 – 08/22 **Birla Institute of Technology and Science, Pilani, Goa, India**
 STUDENT RESEARCHER, Cambridge Centre for Alternative Finance
 Advisors: Snehanshu Saha, Aswini Mishra
 – Developed a self-supervised learning algorithm using a novel reconstruction loss achieving 5% reduction in labelling overhead of continuous tabular data.
 – Conceptualised hybrid GARCH-LSTM ensemble models to predict commodity market volatility and investigated metal market connectedness and volatility spillovers.
 – Investigated the impact of pandemic sentiment on stock market behavior by applying a multi-timescale non-linear causality analysis.

◇ HONORS & AWARDS

- 2018 - 2023 **NATIONAL TALENT SEARCH EXAMINATION (NTSE) SCHOLARSHIP, NCERT**