

Minhwa Lee

minhwalee.ai@gmail.com

[LinkedIn](#)

[Google Scholar Profile](#)

Education

M.S.	Computer Science, University of Massachusetts Amherst, MA USA	09/2021 – 05/2023
B.A.	Computer Science & Mathematics, The College of Wooster, OH USA	08/2017 – 05/2021

Specializations

Natural Language Processing, Human-AI Collaboration, AI for Healthcare

Publications

(Note: Please visit my [Google Scholar](#) profile for the most recent updates. * denotes equal contribution.)

Conference

- 2024 S.A. Hayati, **M. Lee**, D. Rajagopal, and D. Kang. “How Far Can We Extract Diverse Perspectives from Large Language Models?”. **EMNLP 2024**.
- 2024 R. Koo, **M. Lee**, V. Raheja, J.I. Park, Z.M. Kim, and D. Kang. “Benchmarking Cognitive Biases in Large Language Models as Evaluators”. **ACL 2024 Findings**.
- 2024 V. Deshpande, **M. Lee**, Z. Yao, Z. Zhang, J.B. Gibbons, and H. Yu. “LocalTweet to LocalHealth: A Mental Health Surveillance Framework Based on Twitter Data”. **LREC-COLING 2024**.
- 2023 S. Kwon, R. Garodia, **M. Lee**, Z. Yang and H. Yu. “Vision Meets Definitions: Unsupervised Visual Word Sense Disambiguation Incorporating Gloss Information”. **ACL 2023**.

Workshop

- 2024 **M. Lee**, Z.M. Kim, V. Khetan, and D. Kang. “Human-AI Collaborative Taxonomy Construction: A Case Study in Profession-specific Writing Assistants”. **The 3rd In2Writing Workshop @ CHI 2024**.

Preprint

- 2025 L.Wang*, **M.Lee***, R. Volkov, L.T.Chau, D.Kang. “ScholaWrite: A Dataset of End-to-end Scholarly Writing. **Under Review**. 2025.
- 2024 Y.Zhang*, **M.Lee***, J.B.Gibbons, HY Chen, Z. Yao, Y. Wang, O.Bennett, F. Ouyang, D. Levy, K.L.Tucker, and H.Yu. “Socioeconomic and Geographic Disparities in Accessibility to Food Pantries in the United States”. **Under Review**. 2024.
- 2024 D.Das*, K.D. Langis*, A. Martin*, J. Kim*, **M. Lee***, Z.M. Kim*, S.A. Hayati, R. Owan, B. Hu, R.S. Parkar, R. Koo, J.I. Park, A. Tyagi, L. Ferland, S. Roy, V. Liu, and D. Kang. “Under the Surface: Tracking the Artifactuality of LLM-Generated Data”. **arXiv**. 2024.

Professional Experience

2025–Current **Software Engineer @ CitiusTech Inc./Mayo Clinic** *St.Paul, MN (Remote)*

- (1) Working as a senior software engineer (AI Engineering) at Mayo Clinic (Rochester, MN), contracted through CitiusTech US Inc., a leading healthcare IT consulting firm in the US.
- (2) Leading the design and development of an in-house clinical trial matching application that leverages generative AI to extract and analyze patient records for automated eligibility assessment.

- 2024–2025 **Research Assistant @ University of Massachusetts Lowell** *St. Paul, MN (Remote)*
- (1) Fine-tuned several large language models (RoBERTa and LLaMA) on the US Veterans Affairs (VA) electronic health records (EHR) to extract named entities related to social determinants of health (SDOH) and analyzed their statistical association with disease outbreaks among US veterans using statistical regression models.
 - (2) Investigated the influence of implicit biases (e.g., health literacy, educational background, demographic disparities) on the diagnostic accuracy of clinical AI agents in after-visit interactions.
 - (3) Designed an LLM-powered real-time surveillance dashboard for mental health practitioners to assess and monitor suicide risks among US veterans based on clinical notes.
 - (4) Organized and managed university-wide research seminars, streamlining administrative workflows to foster collaboration between faculty and student research assistants.
- 2023–2024 **Visiting Researcher @ University of Minnesota - Twin Cities** *Minneapolis, MN*
- (1) Led an industry collaboration with Accenture Labs to develop a human-AI collaborative prompt curation system for enhancing domain expertise in large language models (LLMs).
 - (2) Constructed a large-scale dataset capturing the human writing process, fine-tuning large language models (LLaMA) with advanced techniques to develop an AI-powered scholarly writing assistant.
 - (3) Designed an open-source, publicly available benchmark dataset for evaluating the reasoning capabilities of large language models under cognitive bias attack scenarios.
- 2023 **Data Scientist Intern @ Microsoft** *Cambridge, MA*
- (1) Designed an end-to-end natural language processing (NLP) pipeline for extracting AI/ML-related keywords automatically from internal project proposals (FY21-22) to reduce human manual workload hours, by fine-tuning NER models (NLTK, SpaCy NER, SciBERT).
 - (2) Developed a real-time Power BI dashboard visualization, integrating insights from fine-tuned language models, significantly improving AI/ML project trend visibility for cross-functional teams.
- 2021–2023 **Graduate Student Researcher @ University of Massachusetts Amherst** *Amherst, MA*
- (1) Developed a multimodal retrieval model, boosting CLIP's image-text matching performance by 10%p using Bayesian inference and GPT-3.
 - (2) Created a large-scale Twitter dataset of mental health and food insecurity, and implemented machine learning models that predict neighborhood-level mental health outcomes from Tweets and CDC survey, using linear regression with pre-trained language models (RoBERTa and Twitter-RoBERTa).
 - (3) Built a web-scraped geolocation dataset of US national food pantries and applied stratified regression models to analyze socioeconomic disparities in food pantry accessibility, visualizing insights in Tableau visualizations for policy recommendations.

Awards & Scholarship

Awards and Honors

- 2021 Finalist, ACM Student Research Competition at Grace Hopper Celebration 2021.
- 2021 Honorable Mention Award, Mathematical Association of America (MAA) Student Poster Presentation in Joint Mathematical Meetings 2021
- 2021 Magna Cum Laude, The College of Wooster
- 2021 Departmental Honors (Computer Science, Mathematics), The College of Wooster
- 2021 The Vivien Chan Prize in Interdisciplinary Sciences, The College of Wooster
- 2019- Pi Mu Epsilon Induction, The College of Wooster

2018-2021 Dean's List, The College of Wooster

Scholarship & Fellowships

2021 Grace Hopper Celebration Student Scholarship

2017-2021 Wooster International Merit Scholarship (\$38,000/yr)

Teaching Experience

Spring 2022 Introduction to Simulation (COMPSCI 550), UMass Amherst

Spring 2021 Introduction to Statistics (DATA 102), The College of Wooster

Fall 2019 Multimedia Computing (CSCI 102), The College of Wooster

Academic Service

2024 – Reviewer, **AMIA Clinical Informatics 2025**

2024 – Reviewer, **ICLR 2025**

2024 – Reviewer, **ACL ARR, EMNLP 2025, ACL 2025, NAACL 2025, ACL 2024, EMNLP 2024**

Updated July 2025