

# Messi H.J. Lee

[messihjee@gmail.com](mailto:messihjee@gmail.com) | [Personal Website](#) | [Google Scholar](#) | [LinkedIn](#)

## Research Interests

---

**Machine Psychology**, AI Bias (e.g., Stereotyping in Large Language and Vision Language Models), Behavioral Evaluation of LLMs, Computational Social Science, Text-as-data

## Experience

---

**MTS Company** | Technical Research Personnel | Republic of Korea June 2025-

- Led the technical aspects of Ministry of Food and Drug Safety approval for pathology AI software;
- Managed IP strategy for Technology-based IPO with 14 patents completed and 22 ongoing.
- Developed AI models to predict lymph node metastasis of breast cancer from core biopsy;

## Education

---

**PhD** | Washington University in St. Louis | United States of America May 2025

- PhD in Computational and Data Sciences (Psychological & Brain Sciences track)
- Dissertation titled, “Stereotyping in Language (Technologies): An Examination of Racial and Gender Stereotypes in Natural Language and Language Models” [[Link to Recording](#)]
- McDonnell International Scholar

**BSc** | Hong Kong University of Science and Technology | Hong Kong 2017-2020

- BSc in Quantitative Social Analysis
- Capstone project titled, “Part-time Electoral Campaigns on Social Media”
- Academic Achievement Medal & First Class Honors

## Full Conference & Journal Publications (Peer Reviewed)

---

Jeon, Soyeon, Lee, M. H. J., Montgomery, J. M., & Lai, C. K. (Political Analysis). From Faces to Politics: Vision-Language Models (Sometimes) Link Visual Demographic Characteristics to Ideological Labels. *Political Analysis*. Published online 2026:1-9. doi:10.1017/pan.2026.10038 [[Link](#)] [[Replication Materials](#)]

Chun, J., Kang, H., Chung, H., Jang, J.-M., Seo, J., Kim, T., Lee, W., Park, C., Hong, M., Kim, H.-M. B., Lee, M. H. J., Jang, K., Jung, C. K., Kim, S. W., & Lee, A. (2026). Deep Learning-Based Diagnosis of Epithelial Ovarian Cancer from Whole-Slide Histopathology Images. *Diagnostics*, 16(10), 1470. <https://doi.org/10.3390/diagnostics16101470> [[Link](#)]

Lee, M. H. J., Montgomery, J. M., & Lai, C. K. (June 2024). Large Language Models Portray Socially Subordinate Groups as More Homogeneous, Consistent with a Bias Observed in Humans. *FAccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*. [[Link](#)] [[Replication Materials](#)]

Lee, M. H. J., Montgomery, J. M., & Lai, C. K. (2024). America's racial framework of superiority and Americanness embedded in natural language. *PNAS Nexus*, pgad485. [[Link](#)] [[Replication Materials](#)]

## Under Review

---

Lee, M. H. J. & Lai C. K. (under review). Implicit Bias-like Patterns in Reasoning Models. [[Pre-print](#)]

PSA005 Team (under review at *Nature Human Behavior*). Stereotype Threat in Black College Students Across Many Operationalizations. [[Stage 1 Report](#)]

Lee, M. H. J., Jeon, S., Montgomery, J. M., & Lai, C. K. (in prep.). Visual Cues of Gender and Race are Associated with Stereotyping in Vision-Language Models. [[Pre-print](#)]

Lee, M. H. J. (under review). Token-Level Entropy Reveals Demographic Disparities in Language Models. [[Pre-print](#)]

## Work in Progress

---

Lee, M. H. J. & Jeon S. (in prep.). Vision-Language Models Generate More Homogeneous Stories for Phenotypically Black Individuals. [[Pre-print](#)]

Lee, M. H. J. & Lai, C. K. (in prep.). Probability of Differentiation Reveals Brittleness of Homogeneity Bias in GPT-4. [[Pre-print](#)]

## Workshop Publications (Lightly Peer-Reviewed)

---

Lee, M. H. J., Montgomery, J. M., & Lai, C. K. (2023). The Effect of Group Status on the Variability of Group Representations in LLM-generated Text. *NeurIPS 2023 Socially Responsible Language Modelling Research (SoLaR) Workshop*. [[OpenReview](#)]

## Talks & Presentations

---

Lee, M. H. J. & Lai, C. K. (Feb 2025). Implicit Bias-like Patterns in Reasoning Models. The Political Data Science Lab Meeting @ Washington University in St. Louis.

Lee, M. H. J., Montgomery, J. M., & Lai, C. K. (Oct 2024). *GPT-4 Associates Distinctively Black Faces with Political Liberalism: Uncovering Political Bias at the Intersection of Race and Gender in AI*. The Political Data Science Lab Meeting @ Washington University in St. Louis.

- Lee, M. H. J., Montgomery, J. M., & Lai, C. K. (June 2024). *Large Language Models Portray Socially Subordinate Groups as More Homogeneous, Consistent with a Bias Observed in Humans*. ACM Conference on Fairness, Accountability, and Transparency 2024 in Rio de Janeiro, Brazil.
- Lee, M. H. J., Montgomery, J. M., & Lai, C. K. (May 2024). *Large Language Models Portray Socially Subordinate Groups as More Homogeneous, Consistent with a Bias Observed in Humans*. Poster Presentation @ STL DataFest 2024 in St. Louis, MO.
- Lee, M. H. J. (Mar 2024). *Large Language Models Portray Socially Subordinate Groups as More Homogeneous, Consistent with a Bias Observed in Humans*. TRIADS “Messy” Data Meeting @ Washington University in St.Louis.
- Lee, M. H. J., Lai, C. K., & Montgomery, J. M. (Feb 2024). *Outgroup Homogeneity-like Effect in Large Language Models*. Poster presentation @ The Society of Personality and Social Psychology Conference in San Diego, CA. [[Poster](#)]
- Lee, M. H. J., Lai, C. K., & Montgomery, J. M. (Dec 2023). *The Effect of Group Status on the Variability of Group Representations in LLM-generated Text*. Poster presentation @ NeurIPS 2023 Socially Responsible Language Modelling Research (SoLaR) Workshop in New Orleans, LA. [[Poster](#)]
- Lee, M. H. J., Montgomery, J. M., & Lai, C. K. (Feb 2023). *A Test of the racial position model in natural language*. Poster presentation @ The Society of Personality and Social Psychology Conference in Atlanta, GA. [[Poster](#)]
- Lee, M. H. J. (April 2022). *Racial/ethnic stereotypes in natural language*. Social and Personality Psychology Brown Bag @ Washington University in St.Louis.
- Lee, M. H. J., & Lai, C. K. (Feb 2022). *The spread of stereotypes on social media*. Virtual poster presentation @ The Society of Personality and Social Psychology Conference.
- Lee, M. H. J., Keller, F., & Yang, J. (Aug 2020). *Central coordination patterns distinguish AstroTurfing from grassroots campaigns*. Virtual poster presentation @ Politics and Computational Social Science & Political Networks.

## Awards

---

<b>Small Grant</b>	2024
The Center for the Study of Race, Ethnicity & Equity at WashU (\$1,000)	
<b>Travel Grant</b>	2024
ACM Conference on Fairness, Accountability, and Transparency (\$1,000)	
<b>OpenAI API Researcher Access Program</b>	2024
OpenAI (\$5,000 in credits)	

<b>Outstanding Research Award</b> The Society for Personality and Social Psychology	2024
<b>Graduate Student Conference Travel Award</b> The Center for the Study of Race, Ethnicity & Equity at WashU (\$750)	2023
<b>Winner and People's Choice of the Digital Transformation Track</b> McDonnell International Scholars Academy (\$3,000)	2023
<b>Small Grant</b> The Center for the Study of Race, Ethnicity & Equity at WashU (\$900)	2022
<b>Diversity Graduate Travel/Registration Award</b> The Society for Personality and Social Psychology (\$500)	2022
<b>Academic Achievement Medal</b> The Hong Kong University of Science and Technology	2020
<b>Undergraduate Research Award Second Runner Up</b> The Hong Kong University of Science and Technology	2020
<b>Scholarship for Continuing Undergraduate Students</b> The Hong Kong University of Science and Technology	2019
<b>Scholarship for Continuing Undergraduate Students</b> Ng Hong Mun Educational Foundation	2017 & 2018

## Teaching

---

<b>Washington University in St.Louis, Assistant in Instruction</b> Text as Data (Instructor: Dr. Brandon Stewart)	2023
<b>Washington University in St.Louis, Assistant in Instruction</b> Quantitative Methods I & II (Instructor: Dr. Shelly Cooper)	2022-2023

## Professional Service

---

### Journal Reviews

- Journal of Language and Social Psychology (May 2025)
- Ethics and Information Technology (Apr 2025)
- Computers in Human Behavior: Artificial Humans (Nov 2024; Jan 2025)
- PLOS ONE (Dec 2024)
- Personality and Social Psychology Bulletin (March 2024)

### Conference Reviews

- Conference on AI, Ethics, and Society (AIES) 2026 (Program Committee)

- 
- Conference on Neural Information Processing Systems (NeurIPS) 2026
  - Conference on AI, Ethics, and Society (AIES) 2025 (Program Committee)
  - Society for Personality and Social Psychology (SPSP) 2024 Student Award
  - Socially Responsible Language Modelling Research (SoLaR) 2.0 Workshop at NeurIPS 2024
  - Association for Computational Linguistics (ACL) Student Research Workshop 2024
  - Trustworthy Multi-modal Foundation Models and AI Agents (TiFA) Workshop at ICML 2024
  - Socially Responsible Language Modelling Research (SoLaR) Workshop at NeurIPS 2023

## Skills

---

### Large Language Models, Natural Language Processing & Text-as-Data

- Retrieval Augmented Generation (RAG), Vector DBs (i.e., ChromaDB)
- Behavioral evaluations of Large Language Models
- Text-as-data approaches (e.g., Structural Topic Models, text classification, sentiment analysis)

### Computational Social Science & Psychology

- Survey experiments using Qualtrics, CloudResearch, MTurk, Prolific

### Programming Languages and Packages Frequently Used

- python: PyTorch, HuggingFace, bitsandbytes, peft, sentence-transformers, transformers, openai, scikit-learn, gensim, numpy, pandas
- R: text, tidyverse, dplyr, ggplot2, psych, lme4, lmerTest, afex, simr

## Research Coverage & Recognition

---

Department of Political Science, Washington University in St.Louis ([Link](#))

2024

Center for the Study of Race, Ethnicity & Equity ([Link](#))