Mao Li

426 Thompson St., Ann Arbor, MI, USA | (734)882-9041 | maolee@umich.edu

SUMMARY

- Expertise in applying Large Language Models to analyze social textual data, with a focus on knowledge distilling and public opinion interpretation.
- Experience in leveraging organic social media data as an alternative to traditional survey methods for trend analysis and public opinion monitoring.
- Contributed to the Census Bureau's data enhancement through stance detection and text summarization, improving data precision and enabling robust comparisons with traditional survey results.
- Collaborated effectively within interdisciplinary teams, encompassing computer science, public health, sociology, and economics, to drive computational social science research, demonstrating strong cross-functional communication and project integration skills.

EDUCATION

Doctor of Philosophy in Survey and Data Science and Scientific Computing (dual degree)University of Michigan, Ann Arbor, MIPresent

Institute for Social Research (ISR) & Michigan Institute for Computational Discovery & Engineering (MICDE)

Master of Science in Survey and Data Science

University of Michigan, Ann Arbor, MI Institute for Social Research (ISR)

Bachelor of Law in Political Science and Administrative Science

Nanjing University (NJU), Nanjing, China School of Government, Department of Political Science

RESEARCH EXPERIENCE

Graduate Student Research Assistant

Project Title: New Approaches to Analyzing Social Media Content for Enhancing Census Bureau Data (PI: Frederick Conrad, Co-PI, Johann Gagnon-Bartsch)

- Developed and applied topic modeling algorithms and text summarization methods to categorize tweets about the Census Bureau and Census 2020 (~10 M), creating a Twitter browser tool that acted as a digital focus group.
- Utilized Large Language Models (LLMs) to accurately detect and classify social media users' stances on various topics. Implemented Parameter-efficient Fine-tuning methods to LLMs for opinion detection contributed to a nuanced understanding of public opinion, aiding in strategy development and informed decision-making.

Dow Distinguished Awardee

Project Title: Pursuing long-term sustainability of green infrastructure from social and economic

- Conducted semi-structured interviews to explore community barriers to adopting sustainable green infrastructure.
- Analyzed qualitative data using text analysis techniques to extract public opinion from interviews, demonstrating proficiency in empirical research and data analysis with real-world applications.

September 2021-Present

July 2019

April 2021

Graduate Student Research Assistant

September 2019-September 2021

Institute for Social Research, University of Michigan

- Project Title: National Neighborhood Data Archive (NaNDA) (PI: Philippa Clarke, PhD)
 Created neighborhood data for the National Neighborhood Data Archive
 - (https://www.openicpsr.org/openicpsr/nanda) at multiple levels of spatial scale across the United States (e.g., Health Care Services, Socioeconomic Status and Demographic Characteristics, Education).
 - Utilized a range of software programs and techniques, including Python, R, ArcGIS and data scraping, to integrate data from multiple sources into a comprehensive suite of neighborhood measures nationally.

Project Title: Cognitive Resilience and Community Context: Examining the role of Neighborhood Built and Social Environments for Slowing the Progression of Dementia among older Americans (PI: Clarke, NIH/NIA 1RF1AG057540)

- Assisted with analyses and writing manuscripts on neighborhoods and cognitive function in a sample of 32,000 aging Americans in the Reasons for Geographic and Racial Differences in Stroke (REGARDS) Study.
- Worked as part of an interdisciplinary team of scientists in public health, sociology, gerontology, psychology, and computer science, to understand the role of neighborhoods for cognitive health.

PEER-REVIEWED MANUSCRIPTS

Li M, & Conrad, F. (2024). Tracking Census Online Self-Completion Using Twitter Posts. *Social Science Computer Review*

Finlay J, Esposito M, **Li M**, Colabianchi N, Zhou HJ, Judd SE, Clarke PJ. Neighborhood Active Aging Infrastructure and Cognitive Health: A Mixed-Methods Study of Aging Americans. *Preventive Medicine*

Finlay J, Esposito M, **Li M**, Khan AM, Gomez-Lopez I, Melendez R, Kobayshi LC, Colabianchi N, Judd SE, Clarke PJ. Can neighborhood social infrastructure modify cognitive function? A mixed-methods study of aging Americans. *Journal of aging and health*.

Finlay J, Esposito M, Yu W, **Li M**, Judd S, Clarke PJ. Brainy neighborhood amenities? A mixed-methods study of intellectually-stimulating infrastructure and cognitive function among older Americans. *Well-being Space and Society*.

Yu W, Esposito M, **Li M**, Judd S, Finlay J. Neighborhood 'Disamenities': local barriers and cognitive function among Black and white aging adults. *BMC public health*.

WORKING PAPERS

Li M, & Conrad F (2024). Advancing Annotation of Stance in Social Media Posts: A Comparative Analysis of Large Language Models and Crowd Sourcing. *arXiv preprint arXiv:2406.07483*.

Li M, Conrad F, & Gagnon-Bartsch J (2024). FastLexRank: Efficient Lexical Ranking for Structuring Social Media Posts. *arXiv preprint arXiv:2410.01183*.

SOFTWARE PACKAGE (Selected)

Twitter Browser (Web App Framework,)

GitHub Repo: https://github.com/tanukinekolee/Tweet_browser

- Project Overview: An in-progress initiative focused on developing a user-friendly web application that provides advanced algorithmic solutions for users without a technical background.
- Contribution: Serving as the lead developer for the implementation of a sophisticated summarization algorithm, enhancing user experience by distilling complex Twitter data into concise, understandable insights.

TECHNICAL SKILLS

- **Programming Language and Software:** Python, C++, R, Spark, Hadoop, Qualtrics, ArcGIS, MySQL
- o Data Collection: Survey, Interviews, and Web Scraping