

Menglin Liu

📍 Davis, CA ✉ mliliu@ucdavis.edu ☎ 530-761-6139 🔗 <https://menglinmileyliu.github.io/index.html>

Experience

Assistant Professor of Computational Social Science

2025 – Present

CUHK-Shenzhen

- Teaching courses on Artificial Intelligence for Social Science and Advanced Statistical Modeling.
- Designing curriculum to integrate applied NLP and machine learning in social science research.

Graduate Researcher

Oct. 2019 – June. 2025

University of California, Davis

- Led development of PoliPrompt, an NLP package integrating prompt optimization and consensus prediction.
- Built production-ready pipelines for LLM-based analysis of survey data and policy meeting transcripts.
- Collaborated with political scientists to ensure interpretability and practical impact.
- Used Weights & Biases, DVC, and Git for experiment tracking and version control.

Data & Policy Analyst

June – Sept. 2023

Committee of 100 (Remote)

- Created interactive geospatial dashboards visualizing property restriction laws.
- Managed end-to-end data pipeline from legislative data ingestion to advanced mapping output.

Professional Summary

NLP Engineer and Computational Social Scientist with Ph.D. in Political Science and 5+ years of experience designing and deploying large-scale NLP pipelines. Proven track record translating political risk and policy questions into production-ready machine learning solutions. Expertise in Hugging Face, PyTorch, generative AI, and interpretability, with open-source contributions and peer-reviewed publications.

Technical Skills

Programming: Python, R, SQL, Stata

NLP / ML Libraries: Hugging Face Transformers, PyTorch, scikit-learn, SpaCy, NLTK, Gensim, TensorFlow

Data and Model Management: DVC, Git, Weights & Biases, MLflow

Cloud / Deployment: AWS (EC2/S3), CI/CD workflows

Other: Causal inference, clustering, regression modeling, survey analysis

Selected Projects

PoliPrompt: LLM-Based NLP Classification Toolkit — Prompt Optimization, Transformers, PyTorch, Open Source

Jun – Sep 2024

- Developed modular NLP pipeline with prompt optimization and exemplar retrieval, achieving 36% accuracy gain and 80% cost reduction.
- Integrated Hugging Face Transformers and PyTorch for scalable training on over 100K text samples.
- Published as open-source PyPI package. [🔗](#) [📦](#)

Scaling Latent Ideology from Congressional Speeches — LLMs, Summarization, Temporal Tracking

May 2025 – Present

- Built NLP pipeline analyzing 20 years of congressional speeches from 200+ members to track ideological shifts over time.
- Used LLM-based summarization and prompting to extract policy-relevant features, supporting pairwise comparisons and Bradley-Terry scaling.
- Produced interpretable temporal ideology scores for political science.

LLM-Based Analysis of Survey Open-Ended Responses — NLP Classification, Prompt Engineering, Topic Modeling, Interpretability

May 2025 – Present

- Designed pipeline classifying 50,000+ open-ended responses for effort, emotion, moral judgment, and ideological stance with LLM prompting.
- Applied topic modeling (LDA) and clustering to uncover latent dimensions of ambivalence, cross-pressure, and cognitive complexity.
- Used interpretable AI techniques with human coding benchmarking to validate theoretical constructs.

Data Processing for Housing Policy Debates — ASR, Diarization, NLP Pipelines, CI/CD

June 2025 – Present

- Automated transcription and speaker diarization for 500+ hours of local government meeting audio.
- Built CI/CD-ready pipelines for text extraction, cleaning, and NLP-based analysis, reducing processing time by 70%.
- Produced structured, quantitative measures of policy discourse for non-technical partners.

Publications and Open Source

- Liu, M. (2025). *Understanding local opposition to housing development: the role of direct democracy in U.S. cities*. [Local Government Studies](#) [🔗](#).
- Liu, M. (2024). *PoliPrompt: A High-Performance Cost-Effective LLM-Based Text Classification Framework for Political Science*. [arXiv preprint](#) [🔗](#).
- **Open Source:** [PoliPrompt on GitHub](#) [🔗](#).

Scholarship and Service

Reviewer for:

- *Social Science Quarterly*
- *Nature Cities*
- *American Journal of Computer Science and Technology*
- *International Communication Association*