**Transparent and Reproducible AI-assisted Annotation for Social Data: An Multi-Agent framework**


- Qunhua Li, Eberly College of Science
- Yuehong (Cassandra) Tai, College of Liberal Arts


Abstract: Transparent, representative, and accountable governance is fundamental to public trust and social justice. Yet analyzing how elected officials communicate with constituents across diverse platforms and contexts poses major methodological challenges. Annotation—a cornerstone of social science research on text and multimedia data—is traditionally slow, costly, and difficult to scale. Large language models (LLMs) offer unprecedented efficiency but introduce new risks: shallow reasoning, misalignment with codebooks, and inconsistent outputs that undermine transparency, reproducibility, and trust in AI-assisted research. This project, led by two PIs with complementary expertise in computational social science and AI for reproducibility, addresses these challenges through two human-centered GenAI frameworks. Aim 1 develops an annotator–inspector dual-agent framework to align LLM reasoning with social science codebooks, producing transparent and reproducible labels of policy domains and stance. Aim 2 introduces a multi-agent, human-in-the-loop system to automate and accelerate codebook refinement, reducing the time and error of traditional adjudication while enabling multimodal annotation. These innovations will be applied to two high-value datasets—both collected and curated by our team: the Digitally Accountable Public Representation (DAPR) Database, containing over 6 million social media posts from 28,000 U.S. officials, and the Local Public Meetings on AI and Governance (LPMAG) dataset, comprising 4,000+ multimodal recordings from 40 states. Applying our frameworks to these data will enable reproducible, large-scale analysis of how public officials communicate, deliberate, and represent their constituents across digital and physical forums, yielding new insights into democratic accountability, participation, and governance. Together, the proposed methods and applications will advance the methodological foundations of responsible, scalable AI in social science, generate open-source tools and datasets for studying democratic communication, and provide hands-on training for graduate and undergraduate researchers—advancing Penn State's strategic goals in interdisciplinary research and student success.