

**Tier 1 Pilot Grant Application: Laying the Foundation
Autumn 2025 Cover Sheet**

Project Information

Project Title	New approach in Nutrition Policy Data Extraction and Categorization: Applying Collaborative LLMs for Cardiovascular Disease Prevention
Budget Request from Initiative	\$25,000
Budget Match (if applicable)	\$0
Total Project Budget	\$25,000

Applicant Information

- Lead Applicant: Yanfang Su, ScD, Assistant Professor

Home School/Department: Department of Global Health, School of Public Health (primary), School of Medicine (joint)

Email: Yfsu@uw.edu

- Co-Applicant: Shuai Huang, PhD, Professor

Home School/Department: College of Engineering, Department of Industrial & Systems Engineering

Email: shuaih@uw.edu

Congjing Zhang, PhD student

Home School/Department: College of Engineering, Department of Industrial & Systems Engineering

Email: congjing@uw.edu

- Finance Point-of-Contact for Lead Department: John Fields jtfields@uw.edu

Abstract

Please provide a description of your project appropriate for a lay reader. Limit your description to no more than 250 words.

We plan to explore a new research direction for an existing policy implementation project funded by NHLBI K01. We have been working on extracting nutrition policies since 2020, and it is labor intensive and Large Language Models (LLMs) might enhance efficiency and scalability in nutrition policy data extraction and categorization. However, LLMs might introduce misinformation through hallucinations and misclassifications especially when ground truth fields are missing or sparse. We propose a Tier-1 pilot to (1) build a benchmarking and evaluation pipeline that distinguishes missing versus misrepresented information (e.g., effective dates, policy types) and (2) prototype a collaborative, multi-agent LLM workflow (e.g., Policy Extractor, Public Health Expert) that iteratively crosschecks outputs for factual consistency and traceable reasoning. The project advances two population-health pillars: (1) human health (by improving quality of public health policy analytics that inform population strategies) and social and economic equity (by increasing transparency of policy data and identifying misinformation patterns that disproportionately affects underserved communities). Over eight months, we will assemble a small, permissioned corpus and demonstrate a working prototype that reduces misinformation relative to LLMs. Deliverables include an open-source benchmark, a modular agent framework, and a preliminary comparative study. These outputs will position our team to pursue a future proof-of-concept proposal to funders (e.g., Knight, Hewlett) focused on trustworthy AI for public health policy.

Tier 1 Project Research Plan

Please provide a narrative description of your proposed project. Your description must include your project aims, description of the problem to be addressed, and your proposed plan, design, and methods. **Limit of one page, including tables and visuals, but excluding references.** References, if needed, should be included as an appendix to the one-page research plan.

Dr. Su has been awarded an NIH K01 to study the impact of public policies on sodium reduction and cardiovascular disease (CVD) prevention. Globally, 3 million people die from excess salt as the leading dietary risk factor for CVD each year and this is a significant public health problem that must be addressed. The U.S. is one of the top three countries with highest population sodium intake per day and one of top four countries with the largest CVD burden; however, there are not much data regarding which sodium reduction policy ‘has been’ adopted in the U.S. We have been working on extracting nutrition policies since 2020, and it is labor intensive and Large Language Models (LLMs) might enhance efficiency and scalability in nutrition policy data extraction and categorization. However, LLMs might introduce misinformation through hallucinations and misclassifications especially when ground truth fields are missing or sparse. We plan to explore a new research direction by exploring the LLMs for an existing policy implementation project funded by NHLBI K01.

This pilot lays the groundwork for a rigorous, completeness-aware evaluation of policy data extraction and a collaborative, multi-agent LLM workflow designed to reduce misinformation. We pursue three tightly coupled aims: (1) construct a benchmark that separates missing values from misrepresented values during evaluation; (2) build a role-specialized, two-agent system (a Policy Extractor and a Public Health Expert) that cross-verifies extractions and rationales; and (3) quantify gains over single-agent LLMs and traditional ML across core tasks, e.g., effective year extraction, policy-type categorization, and factual-consistency checks. Conventional accuracy that treats all blanks as errors can overestimate performance when ground-truth fields are sparse, masking risks from LLM hallucinations and omissions [1, 2]. Health agencies and researchers need extraction pipelines that are both accurate and interpretable, with explicit accounting of missingness and provenance [3, 4, 5]. Our pilot delivers a scalable recipe for trustworthy health policy analytics that directly informs population-health decisions affecting human health and social/economic equity.

		Extraction from LLMs				
		Missing	Present			
Actual information	Missing	Probability = $1 - \alpha$	Probability = α			
	Present	Probability = γ	<table border="1"> <tr> <th>Inaccurate</th> <th>Accurate</th> </tr> <tr> <td>Probability = $\beta - \gamma$</td> <td>Probability = $1 - \beta$</td> </tr> </table>	Inaccurate	Accurate	Probability = $\beta - \gamma$
Inaccurate	Accurate					
Probability = $\beta - \gamma$	Probability = $1 - \beta$					

Notes: The false positive rate is denoted by α , while the overall false negative rate is denoted by β . In our context, any instance where the actual information is present but either missing (γ) or inaccurately extracted ($\beta - \gamma$) by the LLM is classified as a false negative.

This figure introduces our completeness-aware verification framework. When information is present, the model’s inaccurate extractions contribute $\beta - \gamma$; when information is absent, a correct blank contributes $1 - \alpha$. We will report precision/recall/F1 restricted to the non-missing set, together with misinformation rates. This decomposition makes failure modes visible and comparable across tasks and models.

We will implement two LLM agents with complementary roles for policy data extraction. The Policy Extractor LLM performs initial span retrieval and

structured extraction (e.g., dates, policy types) with chain-of-thought hidden but rationales distilled into source justifications. The Public Health Expert LLM critiques outputs for plausibility and consistency with public-health framing (e.g., effective dates, program eligibility). An arbiter routine aggregates the agents’ votes, flags contradictions or low agreement, and triggers minimal human review on high-impact/low-confidence cases. We will ablate prompt templates, retrieval settings, critique styles, and arbitration thresholds.

We will benchmark our proposed method with (1) single-agent LLMs with identical retrieval and prompting, and (2) classic ML/rule-based systems trained on the same corpus. All models will be evaluated under policies with missingness. We will use bootstrap confidence intervals for metric differences, stratify by document type and jurisdiction, and run sensitivity checks for prevalence of missing fields. Error analysis will be organized by failure type ($\alpha, \beta - \gamma, \gamma$).

The project advances two population-health pillars: **human health** (by improving quality of public health policy analytics that inform population strategies) and **social and economic equity** (by increasing transparency of policy data and identifying misinformation patterns that disproportionately affects underserved communities).

References

1. Chen, Canyu, and Kai Shu. "Combating misinformation in the age of llms: Opportunities and challenges." *AI Magazine* 45.3 (2024): 354-368.
2. Chen, Canyu, and Kai Shu. "Can llm-generated misinformation be detected?." arXiv preprint arXiv:2309.13788 (2023).
3. Hart, L. Gary, Eric H. Larson, and Denise M. Lishner. "Rural definitions for health policy and research." *American journal of public health* 95.7 (2005): 1149-1155.
4. Walt, Gill, et al. "'Doing' health policy analysis: methodological and conceptual reflections and challenges." *Health policy and planning* 23.5 (2008): 308-317.
5. Collins, Tea. "Health policy analysis: a simple tool for policy makers." *Public health* 119.3 (2005): 192-196.

Tier 1 Project Evaluation Plan

1. Innovative and Rigorous Metrics

- Deliver a completeness-aware benchmark reporting F1 on non-missing fields, hallucination and omission rates, and provenance-trace scores.
- Target measurable improvements, including 10–15% relative error reduction and reduced hallucination rates compared with baselines.

2. Usability, Transparency, and Open Access

- Build a prototype that provides source-aligned rationales and disagreement flags to enable selective human review.
- Release benchmark code and a reproducible pipeline under a permissive license to ensure transparency and community adoption.

3. Scalability and Future Validation

- Leverage outputs to expand to additional jurisdictions, extend the multi-agent LLM team (e.g., Legal Citations Checker, Provenance Auditor), and conduct prospective validation with policy partners.
- Position the project to secure other funding (e.g., PHI Tier 2, Knight, Hewlett).

Project Timeline

- Month 1: Project launch – begin benchmarking LLMs and misinformation detection framework
- Month 2-3: Literature review and data validation
- Month 4-5: Algorithm development and testing, Mid-term report
- Month 6-8: Evaluation and metrics, draft final report and dissemination plan, project completion

Biographies

Dr. Yanfang Su is an Assistant Professor of Global Health and the Co-founding Director of Learning for Action in Policy Implementation and Health Systems (LAPIS) at University of Washington (UW). Dr. Su is a health economist with an ScD in Global Health and Population from Harvard University. Her research centers on policy implementation science (IS), with a focus on the application of AI in policy analysis. Dr. Su's work has been supported by grants and contracts from federal agencies, foundations, and the private sector. She has authored over 40 peer-reviewed manuscripts, more than 40 non-refereed works, and seven book chapters. As an emerging leader at the intersection of public policy and public health, Dr. Su spearheaded a special issue at *Frontiers* featuring approximately 80 authors, dedicated to leveraging policy implementation insights for global action. Dr. Su has consulted for the Bill & Melinda Gates Foundation, the World Bank, and several NGOs on primary health care system strengthening. She also held academic and policy roles at Tsinghua University and the Hong Kong Policy and Research Institute. Beyond academia, Dr. Su brings over a decade of experience in social entrepreneurship and public health practice, working to strengthen primary health care (PHC) and health policy implementation at global, national, and grassroots levels.

Dr. Shuai Huang is a Professor of Industrial & Systems Engineering at the University of Washington (UW). Shuai's research focuses on developing novel AI and machine learning models for system modeling and decision-making in healthcare and medicine. He received a B.S. in Statistics and a Ph.D. in Industrial Engineering. His major accomplishments include: (1) multi-modal data fusion for precision medicine; (2) AI for mHealth and wearable devices; (3) modeling of networked systems in healthcare and medicine. He has led multi-disciplinary projects funded by NIH, NSF, DARPA, AHRQ, Breakthrough T1D (formerly known as JDRF), Army Research Lab, Air Force Office of Scientific Research, Amazon, Meta, etc.

Tier 1 Project Budget

	Requested from Initiative	Funding Match
Salaries		
Faculty		
Staff	\$15,117	Spatial scientist at CSDE (in-kind)
Student		
Benefits Based on Payroll Load Rate in Effect	\$4,883	
Supplies and Materials Supplies, Equipment Under \$2,000, etc.	-	Computing resources from CSDE (in-kind)
Equipment Equipment Over \$2,000	-	
Tuition		
Other (travel)	\$5,000	
Total Direct Costs (PHI-requested funding cannot exceed \$25K)	\$25,000	

Budget Justification

Briefly detail expenses listed above and explain how funds will be used. Please describe the source(s) of matching funds if included (and include confirmation as an additional page; a .pdf of an email from the individual making the commitment is sufficient). If other sources of funding are being combined with this award to carry out the proposed research, it is critical to address the ways in which the other funding will be used to complete the study or complement the Population Health Initiative award.

Salaries: \$15,117 (28%FTE, 3.35 months)

We will allocate \$15,117 to the research coordinator – corresponding to 3.35 person months.

Supplies, materials, and equipment

We applied to the CSDE in-kind support program, which will allow us to support this budget line if allocated.

Benefits: \$4,883

We allocate 32.3% benefit to the research coordinator as per the FY26 UW benefits rates, representing \$4,883 of the budget

Travel: \$5,000

The travel budget line included airfare tickets, accommodation, travel insurance, and all other travel related expenses of the research and dissemination team.

Funding match (in kind) from CSDE.

Dear CSDE Matching Program Team,

I hope this message finds you well.

I am writing to formally request in-kind matching support from CSDE for a proposal we are preparing for submission to the Population Health Initiative (PHI) Tier 1 Pilot Research Grant. The project is titled:

“New approach in Nutrition Policy Data Extraction and Categorization: Applying Collaborative LLMs for Cardiovascular Disease Prevention”

As part of our interdisciplinary approach, we are seeking the following in-kind resources to strengthen the proposal and ensure successful implementation:

1. **Computing resources** to support regression modeling and data analysis.
2. **Spatial scientist support** to assist with space-time visualization of policy data.

We believe these contributions will significantly enhance the methodological rigor and impact of the project. Please find attached the formal request document outlining the details.

Thank you for considering our application. I would be happy to provide any additional information or meet to discuss further.