Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence (IJCAI-25)
Doctoral Consortium Track

Measuring and Mitigating Homelessness Bias: Leveraging AI for Social Impact

Jonathan A. Karr Jr.*
University of Notre Dame
jkarr@nd.edu

Abstract

Bias towards people experiencing homelessness (PEH) is prevalent in online spaces. I leverage natural language processing (NLP) and large language models (LLMs) to identify, classify, and measure bias using geolocalized data collected from X (formerly Twitter), Reddit, meeting minutes, and news articles across the United States. The results of the study aim to provide a new path to alleviate homelessness by unveiling the intersectional bias that affects PEH. My research delivers a lexicon on homelessness, compiles an annotated database for homelessness and homelessness-racism intersectional (HRI) bias, evaluates LLMs as classifiers against these biases, and audits existing LLMs on HRI. My goal is to contribute to alleviating homelessness by counteracting social stigma and restoring the human dignity of the persons affected.

1 Introduction

During one night in 2024, 771,480 people in the United States were recorded as experiencing homelessness, the highest number ever documented [de Sousa and Henry, 2024]. These numbers are continuing to rise, not only in the United States but throughout the world. The last survey of homelessness from the United Nations found that 100 million people were homeless worldwide [Kothari, 2005]. Specifically, Nigeria, Pakistan, and Afghanistan each have more than 4.5 million people experiencing homelessness (PEH) [World Population Review, 2024]. This shows that the social challenge of homelessness persists throughout the world. Stigmatization of PEH negatively affects the mitigation of homelessness in many ways, since the belief that the poor are undeserving reduces public support for mitigation policies [Applebaum, 2001]. My research aims to contribute to the United Nations' first Sustainable Development Goal: Ending poverty in all its forms everywhere by restoring the agency and capabilities of the PEH [Sen, 2001].

Recognizing the relationship between stigmatization of PEH and efforts to reduce homelessness, I conduct research in natural language processing (NLP) and large language models (LLMs) to identify, classify, measure, and counteract stigma against PEH. My doctoral work aims to answer the following research questions (RQs): **RQ1** - How does homelessness bias vary across US regions, and what factors influence this variation? **RQ2** - How well can existing LLMs classify stigmatization of PEH and HRI bias, and how can their accuracy and performance be improved to meet classification standards? **RQ3** - How biased are existing LLMs towards PEH (auditing)? **RQ4** - How can we counteract homelessness and HRI bias using LLMs and NLP?

2 Methodology

2.1 Prior Work

My Ph.D. work builds on previous research by using NLP and LLMs to identify patterns in bias against PEH by addressing the current limitations of LLMs when it comes to identifying this type of bias, and counteracting it online with LLMs. Existing related work presented an international comparative study on the criminalization of poverty in online public opinion [Curto *et al.*, 2024], and a taxonomy of bias toward aporophobia [Rex *et al.*, 2025]. These studies provided evidence of demeaning attitudes toward PEH [Ranjit *et al.*, 2024]. However, previous studies on homelessness have been limited by lexicons containing a single word, 'homelessness', and by collecting data from a single media source [Ranjit *et al.*, 2024]. Moreover, previous studies do not establish clear correlations that explain the potential beliefs and biases against PEH.

2.2 Project Pipeline

To fill these gaps, I started by creating an initial lexicon with the words: "homeless", "homelessness", "housing crisis", "affordable housing", "unhoused", "houseless", "housing insecurity", "beggar", "squatter", "panhandler", and "soup kitchen". I am using the lexicon to scrape data from small and large counties in the United States that have similar socioeconomic demographic factors, including homelessness rates. To establish the potential association of homelessness bias with racism, I am collecting data in counties with differing racial fractionalization scores to determine whether racial fractionalization contributes to the bias of PEH, as stated in social science literature [Alesina and Glaeser, 2013]. After anonymizing the data, human domain experts will annotate the set of

*Advisor: Nitesh V. Chawla

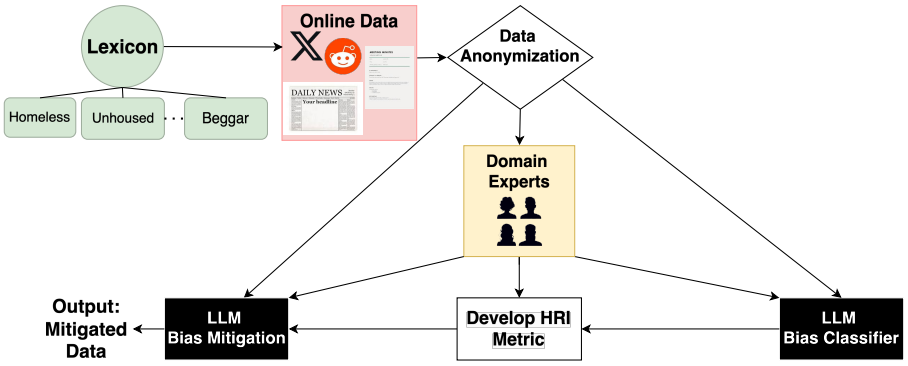

Figure 1: Our pipeline (Section 2.2) begins by using our lexicon to gather online data. We then anonymize it, have experts annotate a portion to create a "Gold Standard," use an LLM to classify the rest, develop an HRI metric, and implement a bias mitigation strategy.

data to define what we consider to be biased. The primary level classifies the text as "direct" or "reporting" bias [Rex *et al.*, 2025] against PEH. This annotated dataset will constitute our "Gold Standard" to determine the efficacy of LLMs in detecting homelessness bias. The second level annotation will be performed based on the topics found by the topic modeling on our initial collected data, similar to previous studies [Ranjit *et al.*, 2024]. Using the results from manual annotation and LLM classifications, I will also develop metrics for HRI bias, using existing lexicons on racism. The results of the LLMs' accuracy in detecting bias will indicate the feasibility of creating an HRI index. The index would allow the tracking of homelessness and HRI bias over time, correlating it with social and political events, to inform homelessness-alleviation policy-making. In addition to evaluating how well existing LLMs can identify bias against PEH and HRI bias, I will evaluate how well LLMs can mitigate it.

3 Limitations and Ethical Considerations

The online data that I am collecting does not cover the entire U.S. population, since 9.78 million people lack broadband access [Palmer, 2025]. I also only use a subset of online resources to analyze English textual data. This does not represent non-English speakers or other communication modalities. Additionally, I am only able to analyze online statements, which may not fully reflect individuals' personal thoughts or feelings. I also recognize the risks and ethical issues inherent in identifying and analyzing negative bias in online spaces [Hovy and Spruit, 2016], including those associated with toxic language detection [Vidgen *et al.*, 2019].

References

[Alesina and Glaeser, 2013] Alberto Alesina and Edward L Glaeser. *Fighting Poverty in the US and Europe*. Oxford University Press, Oxford, 2013.

[Applebaum, 2001] Lauren D Applebaum. The influence of perceived deservingness on policy decisions regarding aid to the poor. *Political psychology*, 22(3):419–442, 2001.

[Curto *et al.*, 2024] Georgina Curto, Svetlana Kiritchenko, Kathleen C Fraser, and Isar Nejadgholi. The crime of being poor: Associations between crime and poverty on social media in eight countries. In *Proceedings of the Sixth Workshop on Natural Language Processing and Computational Social Science (NLP+ CSS 2024)*, pages 32–45, 2024.

[de Sousa and Henry, 2024] Tanya de Sousa and Meghan Henry. The 2024 annual homelessness assessment report (ahar) to congress. Technical report, The U.S. Department of HUD, 2024.

[Hovy and Spruit, 2016] Dirk Hovy and Shannon L Spruit. The social impact of natural language processing. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 591–598, 2016.

[Kothari, 2005] Miloon Kothari. Report of the special rapporteur on adequate housing as a component of the right to an adequate standard of living. Technical report, UN Economic and Social Council Commision on Human Rights, 2005.

[Palmer, 2025] Chili Palmer. Internet facts and statistics, 2025.

[Ranjit *et al.*, 2024] Jaspreet Ranjit, Brihi Joshi, Rebecca Dorn, Laura Petry, Olga Koumoundouros, Jayne Bottarini, Peichen Liu, Eric Rice, and Swabha Swayamdipta. Oathframes: Characterizing online attitudes towards homelessness with llm assistants. *arXiv preprint arXiv:2406.14883*, 2024.

[Rex *et al.*, 2025] Georgina Curto Rex, Svetlana Kiritchenko, Muhammad Hammad Fahim Siddiqui, Isar Nejadgholi, and Kathleen C Fraser. Tackling poverty by acting on social bias against the poor: a taxonomy and dataset on aporophobia. *Forthcoming at the Annual Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics*, 2025.

[Sen, 2001] Amartya Sen. *Development as freedom*. Oxford University Press, 2001.

[Vidgen *et al.*, 2019] Bertie Vidgen, Helen Margetts, and Alex Harris. How much online abuse is there. *Alan Turing Institute*, 11, 2019.

[World Population Review, 2024] World Population Review. Homelessness by country 2024, 2024.