Characterizing Attitudes Towards Homelessness on Social Media

Jaspreet Ranjit
USC Computer Science
jranjit@usc.edu

Rebecca Dorn
USC Computer Science
rdorn@usc.edu

Olga Koumoundouros
USC Suzanne Dworak-Peck
School of Social Work
koumound@usc.edu

Laura Petry
USC Suzanne Dworak-Peck
School of Social Work
lpetry@usc.edu

Eric Rice
USC Suzanne Dworak-Peck
School of Social Work
ericr@usc.edu

Swabha Swayamdipta
USC Computer Science
swabhas@usc.edu

Social media serves as a vehicle for the general public to express opinions and reactions to social events and crises. Complex social issues such as homelessness elicit a diverse spectrum of attitudes, responses and sentiment that differ across socio-political factors such as regionality, ideology and economic climate. The sensitivity and urgency of the crisis people experiencing homelessness (PEH) find themselves in contributes to diverse, nuanced attitudes in the broader society, even as inflammatory comments and harmful biases towards an already marginalized population get the most attention. In either case, there are complex political confounders making it an incredibly challenging task to characterize how people view homelessness. Prior work from the fields of sociology and social work have studied discourse about homelessness grounded in ethnographic studies and qualitative surveys, conducted with few individuals (Tsai et al., 2019; Kingree and Daves, 1997). However, there exists little work to understand societal attitudes about homelessness *at a large scale*, such as expressed in online discourse.

In this extended abstract, we outline an ongoing effort to address this issue by developing structured pragmatic frames to characterize complex societal attitudes towards homelessness across a broad spectrum of social media posts in a U.S. context. Our structured frames are developed using framing theory (Entman, 2006) and codification practices in sociology (Glaser and Strauss, 1967) in collaboration with a team of domain experts from USC Suzanne Dworak-Peck School of Social Work. We ground these frames in data collected from Twitter on the topic of homelessness by creating expert annotations of the posts to characterize them. Figure 1 shows an example of a post annotated with our structured frames: containing coarse and fine-grained frames, as well as additive labels.

However, manual annotation even with domain experts can be laborious and expensive. Given the
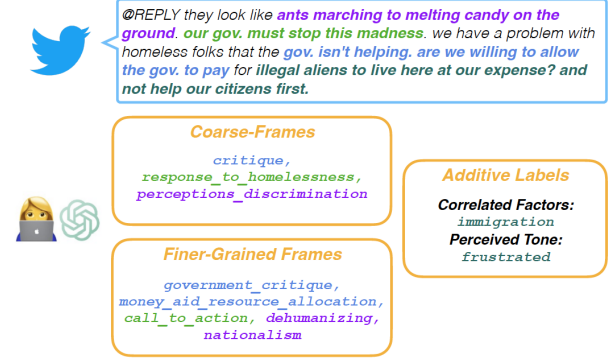

Figure 1: Illustration of proposed annotations for a Twitter post where we use GPT-4 as an assistant in the annotation process to predict coarse frames describing general themes in the post: *coarse_frames*, and further annotate the post with finer-grained frames that fall under the *coarse_frames* and additive labels that contextualize the post with perceived tone, correlated factors, location and subpopulation where applicable.

emergence of large language models in a variety of sociolinguistic tasks (Ziems et al., 2023), we proceed to examine whether these models can be used as effective annotator assistants to the domain experts. Our GPT-4-assisted annotation would involve carefully crafted chain-of-thought prompts (Wei et al., 2022) to ensure the model follows similar reasoning about the nuanced frames as our human experts. However, keeping in mind ethical issues surrounding LLMs, we ensure that our annotations are carefully vetted by our human experts.

Our final collection of data will be used to fine-tune models of natural language to reason about the complex issue of homelessness as well as evaluate them against the expert-curated frames. This will enable us to analyse social media data at scale to automatically infer localized attitudes regarding homelessness as well as analyze changes in attitudes with time. Ultimately, understanding public opinion via online discourse can serve as a useful tool for advocacy groups and political constituents

in tracking attitudes in response to policy and relief efforts to more effectively direct their services in alleviating the suffering of PEH. To summarize, our work focuses on the following research questions:

RQ1 Can we develop general, structured frames on the topic of homelessness, grounded in framing theory, such that these frames can be used to annotate large amounts of data in US context?

RQ2 Using expert-annotated data, can we understand the capabilities and effectiveness of LLMs at reasoning about attitudes towards homelessness?

RQ3 How do socio-cultural, longitudinal, regional and ideological dimensions of frame variation influence attitudes towards homelessness on social media?

1 Curating Online Discourse on Homelessness

Given the diversity (e.g., socio-economic and political backgrounds) of user populations on Twitter[1], it can be understood as an appropriate social media platform for studying different social behavior. Indeed, posts from Twitter have been used to study stigma against compromised populations (Arseniev-Koehler et al., 2016; Kim et al., 2013) and (Mendelsohn et al., 2021) uses framing theory to analyze how immigration is talked about on social media where the results show that adopting frames specific to immigration enabled more obscure yet important information to be captured. Hence, we use online discourse in Twitter posts on homelessness to study and characterize societal attitudes towards homelessness. Specifically, we use a keyword-based search on the Twitter API to randomly sample 3.6 million tweets containing the seed token "homeless".

2 Structured Pragmatic Frames

We target **RQ1** by developing structured pragmatic frames for characterizing attitudes towards homelessness by leveraging framing theory (Entman, 2006). Figure 1 illustrates our current framing typology for a given example post. Our framing typology follows a hierarchical structure where the coarse-frames describing overarching themes that emerged from our qualitative analysis of the tweets including (i) critique, (ii) perceptions and discrimination of PEH, and (iii) responses to homelessness. The finer-grained frames are sub-frames within

each coarse frame that serve as more detailed characterizations of the attitudes towards homelessness in the post. For example, in Figure 1, sub-frames of the critique frame include government_critique and money_aid_resource_allocation about a post discussing government intervention in distributing aid to PEH and immigrants. The additive labels contextualize the tweet with respect to correlated factors, location, subpopulation and perceived tone where applicable. In Figure 1, the relevant additive labels are correlated factors with discussion of immigration policy and perceived tone. Although there is a lot more room for subjectivity when interpreting the tone of a tweet it provides important context on the speaker's attitude with respect to a topic. Offensive tweets occurred frequently in our dataset and perceived tone can serve as a potential indicator for harmful content. We take inspiration from codification in sociology (Glaser and Strauss, 1967) to qualitatively review our tweets and come up with a set of consensus codes that describe attitudes towards homelessness amongst a team of 4 expertly trained annotators. This process was repeated for several pilot studies until we reached thematic saturation at 19 finer-grained labels, grouped across three core frames.

Our work will entail evaluating GPT-4 to serve as an annotator in the loop with human verification to act as an assistant in the annotation process targeting **RQ2**. For evaluating the effectiveness of language models at reasoning about attitudes towards homelessness, we will use our expert-annotated dataset to finetune publicly available models (e.g., FLAN-T5 (Chung et al., 2022)) and automatically infer our frames; we will report metrics corresponding to multi-label classification. We will additionally infer attitudes under a generative setting (by finetuning LLAMA-7B (Touvron et al., 2023)) and evaluate the validity of the generations by human (expert) evaluation. Finally, we plan on answering **RQ3** by conducting analyses with respect to several socio-political dimensions that shed light on how the attitudes towards homelessness change across time spans, by regionality.

We hope that our work can contribute to improved understanding of a dialogue around the crippling issue of homelessness and provide a strong foundation to facilitate meaningful communication between advocacy organizations and community stakeholders.

---

[1]Now, X: https://x.com.

References

Alina Arseniev-Koehler, Hedwig Lee, Tyler H. Mc-Cormick, and Megan Andreas Moreno. 2016. #proana: Pro-eating disorder socialization on twitter. *The Journal of adolescent health : official publication of the Society for Adolescent Medicine*, 58 6:659–64.

Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Eric Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, Albert Webson, Shixiang Shane Gu, Zhuyun Dai, Mirac Suzgun, Xinyun Chen, Aakanksha Chowdhery, Sharan Narang, Gaurav Mishra, Adams Yu, Vincent Zhao, Yanping Huang, Andrew Dai, Hongkun Yu, Slav Petrov, Ed H. Chi, Jeff Dean, Jacob Devlin, Adam Roberts, Denny Zhou, Quoc V. Le, and Jason Wei. 2022. Scaling instruction-finetuned language models.

Robert M. Entman. 2006. Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, 43(4):51–58.

Barney G. Glaser and Anselm L. Strauss. 1967. *Discovery of Grounded Theory Strategies for Qualitative Research*. AldineTransaction, London.

Nathan Kim, Colette Auerswald, and Jessica Lin. 2013. Perceptions regarding homelessness: Analyzing us tweets.

JB Kingree and Walter F Daves. 1997. Preliminary validation of the attitudes toward homelessness inventory. *Journal of community psychology*, 25(3):265–288.

Julia Mendelsohn, Ceren Budak, and David Jurgens. 2021. Modeling framing in immigration discourse on social media. *arXiv preprint arXiv:2104.06443*.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

Jack Tsai, Crystal YS Lee, Jianxun Shen, Steven M Southwick, and Robert H Pietrzak. 2019. Public exposure and attitudes about homelessness. *Journal of Community Psychology*, 47(1):76–92.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35:24824–24837.

Noah Ziems, Wenhao Yu, Zhihan Zhang, and Meng Jiang. 2023. Large language models are built-in autoregressive search engines. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 2666–2678, Toronto, Canada. Association for Computational Linguistics.