

---

# Characterizing Homelessness Discourse on Social Media

**Andrea Hu, Stevie Chancellor, Munmun De Choudhury**

Georgia Tech

Atlanta, GA, USA

{ahu35,schancellor3,munmund}@gatech.edu

## ABSTRACT

Social media allows us to connect and maintain relationships in spite of physical distance and barriers; as computers and the internet become more accessible, hard-to-reach populations are finding a voice on these platforms. One such group is those who are or have been homeless. Through a computational linguistic analysis of a large corpus of Tumblr blog posts, this paper provides preliminary insights to understand the unique ways homeless bloggers express their needs, frustrations and financial/social distress, connect with others, and seek emotional and practical support from others. We highlight future investigations, building upon this research, that can be pursued in HCI to assist an underserved population with the difficult life experience of homelessness.

## INTRODUCTION

For those who are homeless, phones, computers, and social media are incredibly important because they can keep in contact with friends and family while they are transient [7–9, 12, 14]. Different surveys of homeless people in the last decade have found mobile phone ownership ranges from 61.5–76% [7, 8, 12] and that they were not interested in selling their phones to cope with financial hardship—viewing the device as important for them as food [7]. One study done in 2010 found that 48% of their participants used social media such as Facebook [7, 10], while another found that the majority of homeless youth actively maintained their MySpace pages despite internet access challenges [12].

---

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

*CHI'19 Extended Abstracts, May 4–9, 2019, Glasgow, Scotland UK*

© 2019 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-5971-9/19/05.

<https://doi.org/10.1145/3290607.3313057>

**Acknowledgements.** We thank the members of the Social Dynamics and Wellbeing Lab for their valuable feedback during the conceptual stages of the paper. Hu, Chancellor, and De Choudhury were partly supported by a National Institutes of Health grant #R01GM112697.

**KEYWORDS**

homelessness; language analysis; hard to research populations; social media; Tumblr

**Table 1: Example (paraphrased) homeless posts.**

I need money or these people are gonna throw me out on the street again with my 2 dogs. [...] I'm soooo tired!!! And honestly scared. So I haven't made a post like this in forever and I rly don't like doing this but I rly need help!!!

Hi im a 19 year old in <some city> and im living with my mom and our dog in a hotel room, we've been displaced and homeless for a year now and its been extremely hard to get back on our feet. [...] i got stereotyped, accused and fired from my job and still havent been able to pick one up. I know I'm supposed to be reblogging financial help posts, now I'm super pressed to do one myself. Any references, money, help you can offer is appreciated. I live in <some city>; y'all can imagine the huge living expenses here.

**Table 2: Example (paraphrased) not-homeless posts.**

I see homeless people every day. Our lawmakers make legislation that criminalizes being homeless, but being put behind bars repeatedly for small infractions makes it impossible for them to hold down a minimum wage job.

NYC: This whole week is a Code Blue Weather Emergency, meaning that no homeless person seeking shelter can be denied. If you see someone on the street in need of shelter, you can call 311, give a description of the person and the location, and an outreach team will be dispatched to assist. Ok, so one thing I've noticed while working in the food industry [...] People are so mean to homeless people, like ??? There's one homeless guy that walks into my work all the time and all he asks for is a few napkins and to use the bathroom, and my coworkers get mad when I tell him yes

In fact, a study in Seattle in 2011 found 79% of surveyed homeless youth were on social media at least once a week [13]. Consequently, social media applications, like Streetlight Chicago, have been developed to provide information about health care, overnight shelters, drop-in centers and legal aid programs, as well as weather and program closings or changes [11]. Despite challenges of access, it is evident that being homeless does not exclude individuals from using social media, and it can be an important means to connect with peers, family, nonprofit groups, and aid organizations [9, 13].

However, there is an opportunity to understand *how* homeless people use social media. In particular, understanding how homeless people discuss homelessness compared to those who are not can provide us with direct insights into their lives, and how social media can be better tailored to meet their needs. This paper provides preliminary work towards these research investigations.

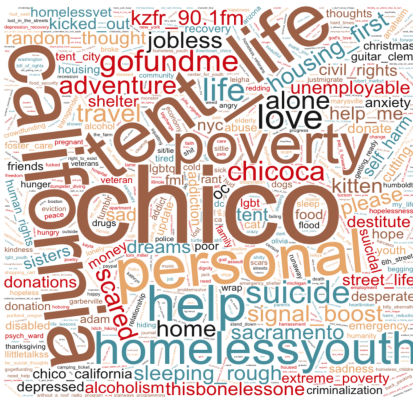
Tumblr provides a unique perspective to examine homeless discourse because it is easy to create pseudonymous accounts on the platform, often used to disclose stigmatizing information [2, 3]. In particular, since Tumblr is a blogging platform, users often write about their life updates and experiences of homelessness in the form of elaborate narratives. Past surveys examining homelessness and social media usage were focused on homeless youth [9, 13], and Tumblr demographics are skewed towards younger populations [1]. Thus, Tumblr provides a suitable platform for understanding the homelessness discourse online—the goal of this preliminary study.

Through a computational analysis of a large corpus of posts on Tumblr made by homeless bloggers, we found significant differences in how homeless and not-homeless Tumblr bloggers appropriated unique combinations of words, topics, and psycholinguistic cues to discuss homelessness. Our results extend prior literature on homelessness and technology-mediated social participation – and also reveal novel opportunities for computational analyses and their outcomes to inform future work to design recommendation and intervention tools for the homeless.

**RELATED WORK**

There have been two approaches towards studying homeless social media use – namely surveys and analyzing social media artifacts. Surveys and interviews have been beneficial because they directly capture the self-reported experiences of homeless people. Social media artifacts have the benefit of being easy to collect and analyze at scale and are less affected by memory and response biases.

Previous work through survey and interview studies on homelessness and social media found that homeless people were utilizing MySpace and, more recently, Facebook [8, 9, 12–14]. In particular, a 2012 study identified Facebook as providing social connections, social support, and a safe environment, but they did not use blogs because they did not think people wanted to read their content [14]. When asked whom they were communicating with on Facebook and MySpace, a study of 23 homeless youth found that most reported using it to keep in touch with people they met on the street, friends from home, and family members; the most commonly reported topics of discussion were music, movies, and videos while shelter was the 7th most commonly reported [13].



**Figure 1: Hashtags used by homeless bloggers.**



**Figure 2: Hashtags used by not-homeless bloggers.**

Previous work analyzing homeless activity on social media has analyzed Twitter. One looked at the social network built from followers of and those followed by WeAreVisible (a homeless visibility project) categorized users into social roles and analyzed the network structure [6]. They found that the homeless individuals connected with WeAreVisible tended to be well-connected with each other, though it was unclear why this was the case or what homeless populations tended to interact with WeAreVisible. Another work studied the values exhibited by homeless and not-homeless individuals over a three-week period using hand-coded tweets [5].

This work aims to expand on our existing understanding of how homeless people use social media by analyzing text blog posts by self-disclosed homeless people and those who have not self-disclosed about homelessness. We hope that studying these artifacts will help elucidate issues around homelessness by highlighting concerns and topics of discussion for an underserved population.

## DATA AND METHODS

**Data Collection and Preparation.** To study homelessness discourse on Tumblr, we collected textual posts related to homelessness from the beginning of the site until March 8, 2018 using the public Tumblr API. We crawled posts with the following hashtags: #homeless, #homelessness, #homeless shelter, #homelessshelter, #poverty, and #begging. This resulted in 26,098 text posts. We chose text posts because they can give insight into discussions surrounding homelessness and can be analyzed in a scalable and automated way using computational language analysis techniques.

Next, we identified posts made by homeless people – our corpus likely contains noise, such as nonprofits or advocates using social media to raise awareness. Two researchers working in the field of social computing labeled 1,000 randomly sampled posts to determine whether they contained a self-disclosure of being homeless or reported having been or about to be homeless recently. All other posts were marked as non-self-disclosures. Raters achieved an interrater agreement score of Cohen’s  $\kappa = 0.78$ . Examples of these two types of posts are given in Tables 1 and 2.

To expand the dataset and drawing on prior work [4], we assumed that a blog with a post marked as a self-disclosure of homelessness is owned by a homeless person; therefore, all of their posts about homelessness are from the perspective of a homeless person. Next, we looked at blogs that created both self-disclosure and non-self-disclosure posts and manually examined the veracity of our assumption by looking at all posts they had made and determining their homeless status. We define *homeless bloggers* as those that had written a self-disclosing post that they had been homeless in the labeled dataset, while *not-homeless bloggers* did not write any posts self-disclosing that they were homeless in the labeled dataset. All posts (regardless of whether they contain an explicit self-disclosure) that were written by homeless bloggers are referred to as *homeless posts* and all posts written by not-homeless bloggers are referred to as *not-homeless posts*. Before expanding the dataset, we had 179 self-disclosure posts and 821 not-self-disclosure posts. From these posts we determined there were 142 homeless bloggers and 739 not-homeless bloggers in the labeled dataset. After using the bloggers' homeless status to expand the dataset, we had 580 homeless posts and 1,996 not-homeless posts.

<sup>1</sup><http://liwc.wpengine.com/>

<sup>2</sup><https://radimrehurek.com/gensim/models/word2vec.html>

**Table 3: Word embeddings (or most similar words) trained on homeless blog posts corresponding to the top five most common hashtags in the dataset.**

Hashtag	Most similar words
homeless	tent, night, new, day, love, living, city, shelter, people, good, food, car, like, community, feel, chico, roof, thing, job, today
help	time, need, little, money, like, said, good know, things, think, car, make, work, way, maybe, place, didn, thing, want, right
life	homelessness, tent, homeless, city, kitten, shelter, day, living, god, new, cat, tonight, love, mental, want, real, year, chico, night, street
food	like, need, way, little, car, time, people, money, good, went, right, night, things, thing, feel, them, work, house, maybe, make

**Table 4: Word embeddings (or most similar words) trained on not-homeless blog posts corresponding to the top five most common hashtags in the dataset.**

Hashtag	Most similar words
homeless	compassion, poverty, criminalization, charity, info, njcounts, anne, naeh, inspirational, cbpp, causes, alliance, allison, belfast, notforprofit, end, lgbtq, hiv, interagency, samaritan
help	person, come, work, friends, job, this, place, talk, away, family, different, world, live, water, try, girl, instead, fucking, getting, sure
life	sign, look, them, all, seen, different, change, place, this, walk, home, come, took, live, sleep, think, asking, try, hope, start
food	people, guy, begging, hungry, night, person, guess, help, jesus, donate, visit, places, turn, family, water, came, man, couple, world, true

**Computational Approach.** When analyzing the text of these blogs, we considered text to be a combination of the blog title, text, and the hashtags. We used established techniques: Linguistic Inquiry and Word Count (LIWC)<sup>1</sup>, hashtag comparisons using log likelihood ratios, word embeddings (specifically the Word2Vec algorithm<sup>2</sup>) and Latent Dirichlet Allocation (LDA) based topic modeling to analyze the differences in the ways the two groups discuss homelessness. These different techniques can provide a multi-faceted view of homelessness discourse on Tumblr: LIWC provides insights about the psycholinguistic use of hand-curated categories of words, hashtag use demonstrates information dissemination behaviors, word embeddings highlight the linguistic context of use of specific words, and LDA helps identify topics and themes in text in an unsupervised way.

## RESULTS

*Psycholinguistic Analysis.* Beginning with an analysis of psycholinguistic word use, we find a number of word categories that were statistically significantly more widely used by the homeless group. Most importantly, emotion words belonging to the categories negative affect (NA), anger, sadness, and anxiety were significantly higher in the homeless group by 12-54% compared to the other (Mann-Whitney U tests gave significance at the  $p = 0.0001$  level, following Bonferroni correction). Additionally, social and personal concern related words such as those about ‘work’, ‘social’, ‘money’, ‘friends’, ‘health’, ‘body’, and ‘humans’ were also higher by 16-44% ( $p < 10^{-8}$ ). A greater self-attentional focus was apparent in this group via the use of greater number of first person pronouns (39% higher). Together, these indicate that discourse for homeless bloggers is largely about negative personal experiences and struggles with aspects of daily life. Note that familial words, unlike friend words, were *not* discussed in much elaboration in the homeless blogs, even though it is often these reasons why homeless people go out of their way to get access to phones, computers, and social media [7–9, 12, 14]. It is possible that differences in target audience on Tumblr might account for why friend related words are more common versus family ones.

*Hashtag Use.* Next, we present hashtags used by the two groups as word clouds (where size of a word corresponds to the number of occurrences) to examine their differences in disseminating information on Tumblr and engaging with the broader social audience on the platform. Figure 1 shows the hashtags that were more frequently used by the homeless bloggers, based on the log likelihood ratio (LLR) analysis, that computes the logarithm of the ratio of the probability of occurrence of a word in one sample to that in a second sample. We observe the presence of words relating to financial hardships (‘poverty’, ‘destitute’, ‘poor’), those about mental and behavioral health issues (‘suicide’, ‘anxiety’, ‘alcoholism’, ‘self\_harm’), professional distress (‘jobless’, ‘unemployable’), and practical challenges of homeless living conditions (‘housing’, ‘tent\_life’, ‘hunger’). On the other hand, the non-homeless bloggers (ref. Figure 2) discusses charity, and general discourse about helping homeless people via advocacy and acts of kindness (‘non-profit’, ‘charity’, ‘kindness’, ‘gratitude’, ‘generosity’).

**Table 5: LDA topics from homeless blog posts.**

Topic	Words
0	kitten, cat, like, tent, trying, big, maybe, today, homelessness, won
1	homeless, secretly, thank, christmas, car, help, mom, little, stop, homelessness
2	homeless, people, homelessness, like, know, time, life, chico, help, day
3	housing, chico, sitlie, homeless, laws, action, radio, butte, julie, tiny
4	family, homeless, like, new, home, moved, parents, away, lived, right
5	billy, hey, says, like, return, blanket, thing, homeless, things, people
6	people, place, homeless, life, said, help, tent, hope, dad, asking
7	dog, mama, pet, rose, pets, dinner, sacramento, dogs, project, thanksgiving
8	veterans, stand, housing, veteran, hud, assistance, tom, homeless, california, words
9	tent, day, life, like, homelessness, little, night, unemployed, think, morning
10	love, mother, told, keys, called, appreciate, friends, wrong, said, homeless
11	rights, kayla, sanfrancisco, sabbath, paul, god, boden, homeless, faith, wrap
12	paul, boden, advocacy, regional, project, western, wrap, radio, roof, kzfr

*Word Embeddings and Topics.* Third, we discuss the context of use of some of these words used by the two groups by looking at their embeddings in a higher dimensional space. Specifically, the Word2Vec algorithm uses cosine similarities to discover the most similarly used words given a word in the corpus. Using this technique, we find the top 20 words most similar to the most commonly occurring hashtags obtained above, “homeless”, “help”, “life”, “food”, and “poor”: this is given in Table 3 for the homeless bloggers and in Table 3 for the not-homeless bloggers. We observe that homeless bloggers tend to use words describing homeless life, while non-homeless bloggers use more abstract words to describe and discuss homelessness, such as ‘poverty’. An interesting observation is that “car” appears prominently in the top 20 most similar terms for all five hashtags (except for “life”) when trained on the homeless blog posts. This suggests that cars are important to homeless bloggers. An inspection of posts containing this word describe it as a temporary shelter and one of the few things the homeless personally own, even after losing a place to live.

Finally, LDA produced 13 topics in the homeless corpus, shown in Table 5, while we obtained 17 topics from the posts in the non-homeless dataset, which are shown in Table 6. Aligning with the above insights, we observe that both groups discuss aspects related to homeless advocacy and support organizations for the homeless. We also find that homeless bloggers discuss topics from their own life more frequently while not-homeless bloggers discuss the news more.

## DISCUSSION, CONCLUSION, AND FUTURE WORK

**Main Insights.** Aligning with prior work that qualitatively analyzed Twitter posts of the homeless [5], we found significant frequency differences in how homeless and not-homeless bloggers discussed various issues, such as community support, wealth, and economic prosperity. Drawing upon Eyrich-Garg’s value framework concept of social connectedness [5], it is apparent that the homeless bloggers used Tumblr for staying socially connected (with friends and peers), while navigating the challenging conditions of homelessness. We also found Tumblr to provide a unique mechanism for homeless bloggers to seek financial help from private donors as well as charity organizations and nonprofits. As Andalibi et al. notes [2], Tumblr use thus helped these bloggers meet both their emotional needs (e.g., venting out frustration, seeking support, finding and maintaining dignity, self-esteem, resilience, and psychological wellbeing surrounding a stigmatizing experience), as well as practical ones (e.g., gather pointers to financial, shelter, and employment resources). At the same time, that homelessness is discussed beyond the personal experiences of this condition complementarily indicates the use of Tumblr as a platform for raising awareness on this societal issue, for spreading advocacy, for discussing affected individuals’ civil rights and legal and law enforcement dimensions, and for finding a means to connect with otherwise underserved and hard to reach populations.

**Recommendations for Future Investigations.** Our results open up many opportunities for future investigations. An important novel direction includes HCI research targeted at designing social media sites and applications to be more effective at meeting homeless individuals’ needs. Since our research

**Table 6: LDA topics from non-homeless blog posts.**

Topic	Words
0	francis, pope, square, sleeping, bags, open, visit, service, free, news
1	homelessness, housing, new, veterans, services, end, families, view, people, program
2	volunteers, eat, serving, food, opportunities, love, time, room, help, nice
3	art, photography, book, images, faces, rich, middle, life, piece, phone
4	homeless, people, like, man, said, know, money, food, time, day
5	dog, dogs, animals, begging, america, homeless, post, hospital, car, funny
6	god, jesus, church, love, homeless, people, life, lord, christ, like
7	man, police, view, video, woman, officer, bus, cops, law, officers
8	people, new, home, help, children, homelessness, shelter, family, years, city
9	beach, hours, seen, hour, best, begging, husband, holiday, cold, tiny
10	pretty, fun, amazing, sun, information, media, cool, guess, sunday, piece
11	youth, lgbtq, gay, like, foster, said, time, home, house, eyes
12	music, looks, people, moving, like, old, remember, noticed, felt, section
13	kindness, poverty, homeless, view, wordpress, acts, photo, homelessness, food, love
14	van, die, met, door, word, man, open, homeless, tomorrow, winter
15	albuquerque, worst, case, lived, second, clients, men, today, start, stand
16	alliance, esg, rapid, rehousing, webinar, prevention, hprp, hud, click, program

leverages established computational techniques to understand and identify these needs, it provides a scalable, and semi-automated way to surface these needs to appropriate and interested parties, such as non-profits and homeless advocacy organizations. It is also possible to consider building near real-time recommendation and intervention tools for homeless social media users—e.g., their needs and calls for help could be followed up with pointers to local resources, such as homeless shelters or short-term employment opportunities. We found homeless bloggers already using sites like GoFundMe and PayPal to collect online donations; these adhoc practices can be made more streamlined and seamless within the social media platforms to connect those in need with donors. Finally, conducting follow-up interviews and/or surveys to go alongside this type of computational analysis will provide opportunities to compare expressed needs and challenges with perceived ones and to explore the benefits (or not) of social media as a means of social support and group advocacy.

## REFERENCES

- [1] 2015. Demographics of Tumblr. [http://www.pewinternet.org/2015/08/19/mobile-messaging-and-social-media-2015/2015-08-19\\_social-media-update\\_05/](http://www.pewinternet.org/2015/08/19/mobile-messaging-and-social-media-2015/2015-08-19_social-media-update_05/)
- [2] Nazanin Andalibi, Margaret E Morris, and Andrea Forte. 2018. Testing waters, sending clues: Indirect disclosures of socially stigmatized experiences on social media. *Proc. ACM - HCI 2, CSCW* (2018), 19.
- [3] Munmun De Choudhury and Sushovan De. 2014. Mental Health Discourse on reddit: Self-Disclosure, Social Support, and Anonymity.. In *ICWSM*.
- [4] Munmun De Choudhury, Sanket S Sharma, Tomaz Logar, Wouter Eekhout, and René Clausen Nielsen. 2017. Gender and cross-cultural differences in social media disclosures of mental illness. In *Proc. CSCW. ACM*, 353–369.
- [5] Jes A Koepfler and Kenneth R Fleischmann. 2012. Studying the values of hard-to-reach populations: content analysis of tweets by the 21st century homeless. In *Proceedings of the 2012 iConference. ACM*, 48–55.
- [6] Jes A Koepfler and Derek L Hansen. 2012. We Are Visible: Technology-mediated social participation in a Twitter network for the homeless. In *Proceedings of the 2012 iConference. ACM*, 492–493.
- [7] Christopher A Le Dantec and W Keith Edwards. 2008. Designs on dignity: perceptions of technology among the homeless. In *Proc. CHI. ACM*, 627–636.
- [8] Christopher A Le Dantec, Robert G Farrell, Jim E Christensen, Mark Bailey, Jason B Ellis, Wendy A Kellogg, and W Keith Edwards. 2011. Publics in practice: Ubiquitous computing at a shelter for homeless mothers. In *Proc. CHI. ACM*, 1687–1696.
- [9] Eric Rice and Anamika Barman-Adhikari. 2014. Internet and social media use as a resource among homeless youth. *Journal of Computer-Mediated Communication* 19, 2 (2014), 232–247.
- [10] Eric Rice, Alex Lee, and Sean Taitt. 2011. Cell phone use among homeless youth: potential for new health interventions and research. *Journal of Urban Health* 88, 6 (2011), 1175–1182.
- [11] Haley Velasco. 2017. How Homeless Chicago Teens Are Using Social Media to Communicate. <https://news.medill.northwestern.edu/chicago/how-homeless-chicago-teens-are-using-social-media-to-communicate/>
- [12] Jill Palzkill Woelfer and David G Hendry. 2010. Homeless young people's experiences with information systems: life and work in a community technology center. In *Proc. CHI. ACM*, 1291–1300.
- [13] Jill Palzkill Woelfer and David G. Hendry. 2012. Homeless Young People on Social Network Sites. In *Proc. CHI (CHI '12). ACM*, New York, NY, USA, 2825–2834. <https://doi.org/10.1145/2207676.2208686>
- [14] Mary Yost. 2012. The invisible become visible: An analysis of how people experiencing homelessness use social media. *Elon Journal of Undergraduate Research in Communications* 3, 2 (2012).