Under the (Neighbor)hood: Hyperlocal Surveillance on Nextdoor

Madiha Z. Choksi*
mc2376@cornell.edu
Cornell Tech
New York, New York, USA

Marianne Aubin Le Quéré*
msa258@cornell.edu
Cornell Tech
New York, New York, USA

Travis Lloyd
tgl33@cornell.edu
Cornell Tech
New York, New York, USA

Ruojia Tao
rt456@cornell.edu
Cornell Tech
New York, New York, USA

James Grimmelman
james.grimmelmann@cornell.edu
Cornell Tech
New York, New York, USA

Mor Naaman
mor.naaman@cornell.edu
Cornell Tech
New York, New York, USA

ABSTRACT

This paper examines the tensions between neighborhood gentrification and community surveillance posts on Nextdoor, a hyperlocal social media platform for neighborhoods. We created a privacy-preserving pipeline to gather research data from public Nextdoor posts in Atlanta, Georgia and filtered these to a dataset of 1,537 community surveillance posts. We developed a qualitative codebook to label observed patterns of community surveillance, and deploy a large language model to tag these posts at scale. Ultimately, we present an extensible and empirically-tested typology of the modes of community surveillance that occur on hyperlocal platforms. We find a complex relationship between community surveillance posts and neighborhood gentrification, which indicates that publicly disclosing information about perceived outsiders, especially for petty crimes, is most prevalent in gentrifying neighborhoods. Our empirical evidence inform critical perspectives which posit that community surveillance on platforms like Nextdoor can exclude and marginalize minoritized populations, particularly in gentrifying neighborhoods. Our findings carry broader implications for hyperlocal social platforms and their potential to amplify and exacerbate social tensions and exclusion.

CCS CONCEPTS

• **Security and privacy → Social aspects of security and privacy**; • **Human-centered computing → Empirical studies in collaborative and social computing**; **Social media**.

KEYWORDS

Privacy, Social Media, Online Communities, City, Method, Theory, Qualitative Methods, Quantitative Methods

ACM Reference Format:
Madiha Z. Choksi, Marianne Aubin Le Quéré, Travis Lloyd, Ruojia Tao, James Grimmelman, and Mor Naaman. 2024. Under the (Neighbor)hood: Hyperlocal Surveillance on Nextdoor. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 22 pages. https://doi.org/10.1145/3613904.3641967

*Both authors contributed equally to this research.

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3641967

1 INTRODUCTION

Hyperlocal social platforms serve as vibrant hubs for local information exchange and play a crucial role in cultivating community engagement [23, 61, 66, 72]. Despite their widespread adoption and appeal as virtual town squares, location-based social platforms have increasingly come under scrutiny for the real-world harms these often unmonitored spaces perpetuate, particularly for minoritized communities.

The tension between helpful neighborhood exchange forum and toxic social platform is especially pronounced in discussions concerning Nextdoor [36]. As of 2022, Nextdoor is the largest local social media platform, hosting communities in over 180,000 neighborhoods across the United States [2]. Critics have observed that the neighborly façade of Nextdoor may conceal racist and toxic content posted under the guise of "civility" [36, 37, 48, 78]. The media has also highlighted instances of user posts on Nextdoor which fixate on "suspicious-activity" carried out by *others or outsiders* that disproportionately target people of color [78]. In Oakland, for example, Levin [52] documented racial profiling of Black residents by their white neighbors [52]. Other work has demonstrated that Nextdoor posts can amplify community surveillance to produce harms such as social exclusion and gentrification [47, 49]. While empirical findings have largely confirmed these critiques, this growing body of work on hyperlocal social platforms has yet to trace the qualitative dimensions of *how* such harms are produced in the language, style, and patterns of user-generated posts across communities. We fill this gap with an empirical study exploring the patterns of community surveillance content on Nextdoor.

User-generated community posts – primarily concerned with neighborhood crime and safety – are examples of *community surveillance*, which is defined as the systematic monitoring and observation of activities, behaviors, or information within a specific community or neighborhood by its own members [45]. This form of surveillance is motivated by a community's collective interest in gathering information related to their own well-being by promoting and maintaining community safety [15]. Community-surveillance strategies often involve the collection and analysis of local information on a voluntary or informal basis to help community members make informed decisions or address concerns related to their local environment, social interactions, or public safety [71].

Examining crime- and safety-related content on Nextdoor provides an opportunity to understand how user-generated content contributes to modern community surveillance. In 2013, crime- and safety-related content accounted for approximately 20% of all content on Nextdoor [48]. In the first large-scale study of Nextdoor data, Iqbal et al. categorized Nextdoor crime and safety posts broadly into “Drugs and Order,” “Theft and Property Damage,” and “Weapons and Violent Crimes,” demonstrating that these types of posts persist on the platform. This work also shows that correlations between official crime statistics and discussions of crime on Nextdoor are influenced by neighborhood income, where higher-income neighborhoods are more proportionally likely to discuss crime on the platform [38, 39]. Similarly, in their analysis of Nextdoor posts in Philadelphia, Lee and Ahn find that while open discrimination by race on Nextdoor is infrequent, posters frequently employ postracial strategies to implicitly discriminate against minorities [51].

At a high level, our study bridges existing empirical work on Nextdoor that captures high level patterns of community surveillance via user-generated content (posts) with longstanding work in the surveillance studies literature. We do so by contributing a fine grained qualitative analysis of 1,537 community surveillance posts. Our study complements existing high-level statistical analyses through a qualitative labeling exercise of community surveillance posts, thus allowing us to understand how online content contributes to larger trends. To that end, we ask:

- **RQ1:** What are the patterns of community surveillance in user-generated posts within Nextdoor neighborhoods?
- **RQ2**: Is there a relationship between neighborhood gentrification trends and types of community surveillance posts?

We address these questions and contribute a more nuanced understanding of Nextdoor's impact by creating a typology of community surveillance posts, and correlating these groupings with neighborhood gentrification trends. Our study introduces a mixed-methods analysis of posts on Nextdoor in Atlanta, Georgia, USA. We construct a privacy-preserving, neighborhood-level pipeline to collect public Nextdoor posts, allowing us to gain access to novel research data. We filter for posts that match a set of community-surveillance keywords, which we manually validate, resulting in 1,537 relevant posts for further analysis. We develop and apply a qualitative codebook to each Nextdoor post that includes information about the poster, the post subject, the communication tactic used in the post, and the emotion that accompanied the post. We develop and validate a novel approach to using a large language model (LLM) to label qualitative data, contributing an extensible framework to tag community surveillance posts on social media. Our analysis visualizes these codes using Multiple Correspondence Analysis (MCA) to demonstrate how each of these dimensions relate to neighborhood gentrification, and elaborate on how gentrification trends inform communication tactics. We further analyze the potential harms of each type of community surveillance post given Nextdoor's known relationship with the police and government [36]. Finally, we theorize how specific types of community surveillance in hyperlocal digital platforms become a collective tool used to exclude minoritized populations in gentrifying neighborhoods.

2 RELATED WORK

The theoretical backdrop for our research are frameworks that discuss the relationships between privacy, digital connectivity, and surveillance [12, 29]. These frameworks serve as a critical lens through which we construct our study design to qualitatively assess surveillance patterns in user-generated content. We map these ideas to the broader context of Human-Computer Interaction (HCI) studies on social platforms and summarize empirical studies on Nextdoor which examine community dynamics and surveillance harms. These prior works uncover how social platforms like Nextdoor facilitate community engagement while monitoring neighborhood activities, and unveil the intricate relationship between platforms, socioeconomic changes, and neighborhood transformations [25, 26, 47].

2.1 Privacy and Surveillance Theory

Our work draws on theoretical conceptualizations of surveillance as *“the pervasive monitoring of individuals or groups”* [29]. According to Foucault, societies devise institutional surveillance mechanisms to control and normalize behavior *en masse* by integrating an ominous fear of discipline [29, 57, 58]. In digital spaces, notions of surveillance extend into networked infrastructures, where platforms and technologies facilitate further forms of economic surveillance, or surveillance capitalism [12, 81]. These surveillance mechanisms serve to *classify* individuals, aligning their online behavior with the interests of political actors involved in their development and implementation [81].

As surveillance methods become increasingly accessible, automated, and cheap, ubiquitous surveillance practices shift towards influencing behavior through the systematic awareness of people's lives [81]. Information-based surveillance practices permeate institutional routines and the surveillance of everyday life ultimately becomes a central feature of modern social relations. In the case of Nextdoor, this manifests a crowd-sourced surveillance ecosystem that creates a form of social control over users' behavior. Further, Nextdoor's platform affordances incentivize oversharing, thus normatively compelling and rewarding community members who conform to certain behaviors [7, 59].

Scholars highlight that tenets of privacy are often in tension with the stated objectives of social platforms. Such platforms create a delicate balance between a user's desire for privacy and their need for social connections [7] and this has been observed across popular social platforms that pre-date Nextdoor. Facebook, oriented towards connecting individuals, compels users to disclose what Facebook defines as a “real name” under a strict “authentic identities” policy [32]. On Facebook, users voluntarily share personal information, effectively blurring the lines between their online presence and offline identity [32]. In contrast, Reddit thrives on generating and aggregating content for discussions, allowing users to post under pseudonyms [27] to maintain safe yet candid spaces for community interactions. Despite their divergent approaches to privacy and anonymity, both platforms align these strategies with their understanding of hosting a thriving community [14].

In the context of HCI, Nextdoor emerges as a unique case study to explore digital privacy tensions. Nextdoor constructs a vision of community through shared geographic proximity in well-defined "neighborhoods," grouping people together on that basis. Its design fundamentally hinges on information disclosure (e.g. real name and physical address), and the absence of anonymity and privacy [62]. Given Nextdoor's unique geographic orientation, privacy norms on Nextdoor do not map neatly onto other social platforms. The distinction is that the tangible, real-world connections among users—these are not just virtual profiles but real people, real neighbors. As a result, the dynamics of privacy on Nextdoor are intricately tied to the dynamics of physical community life, introducing a novel dimension to the discourse on digital privacy. Exploring these distinctions provides valuable insights into the evolving landscape of online interactions within the context of genuine, offline communities.

2.2 Community Policing, Harms, and Gentrification

Several studies on community policing technologies have focused on crime prevention [8, 13, 18, 53]. Lewis and Lewis [53] illustrate how residents use a crime-prevention forum to forge social connections, organize collective actions, exchange information, and establish and enforce social norms both within the neighborhood and on the web forum itself [53]. Ceccato [13] find that crime-prevention apps are primarily used to reactively report incidents rather than to proactively prevent them. Popular safety applications frame crime as endemic in urban spaces, perpetuating a false sense of vulnerability while ignoring the structural factors that underlie criminal activity [44]. Unpaid and untrained neighborhood moderators, coupled with algorithms designed to maximize user engagement, can fuel polarization and contribute to the amplification of controversial subjects [44].

Researchers have also theorized how technically constructed neighborhoods on Nextdoor perpetuate offline inequality and the exclusion of minority groups [8, 47, 49, 65]. Payne [65] contends that defining Nextdoor neighborhoods in a polygonal way with rigidly defined borders cleaves communities along socioeconomic lines [65]. Specifically, spatial fragmentation is most visible in high-rent districts where wealthier residents create fewer, smaller neighborhoods [65]. Similarly, Kurwa [47] observes that Nextdoor builds digitally gated communities, replicating the exclusions historically associated with physical gated communities. Divisions between neighborhoods has particularly strong implications for neighborhoods trending towards integration (when minorities move into primarily white neighborhoods) and gentrification (when white individuals move into minority neighborhoods) [47].

Lambright [49] compares Nextdoor's stated community values to "the midcentury imagination of what a community should be," arguing that it is unsurprising that the platform has also replicated a space for "digital redlining." Finally, Bloch [8] theorizes that the unconscious exclusion and avoidance of black people by white people on Nextdoor platform creates a space for aversive racism to proliferate.

Empirical explorations of user-generated Nextdoor content have started to corroborate theoretical arguments. Iqbal et al. [38] report a strong correlation between economic disparities and user-generated discourse on Nextdoor where wealthier neighborhoods exhibit more positive sentiment and discuss crime proportionally more frequently, despite having lower crime rates. Lee and Ahn [51] find that Nextdoor users employ subtler strategies to normalize and conceal anti-Black sentiments and settler ideologies within the broader socio-cultural and political context. These strategies embed racialization in policy discussions while avoiding explicitly problematic language. These studies reinforce the need to qualitatively assess the semantic styles and patterns of user-posts and their relationship with gentrification trends on Nextdoor.

The human-computer interaction field (HCI) has also engaged with Nextdoor specifically and neighborhood safety apps more broadly. In an empirical multi-neighborhood community-based study of Nextdoor in Atlanta, Masden et al. [59] find that the absence of anonymity has the potential to generate harms. While there was a positive correlation between community engagement on civic issues, the majority of the study's participants expressed major privacy concerns such as how much information could be gleaned about their daily habits. To mitigate harms, residents chose to self-moderate their content, revealing and concealing bits of information and even removing information that might compromise the privacy of their home-life. An analysis of Citizen, another popular neighborhood safety application, found that deceptive design patterns that promote and capitalize on neighborhood fear can be built into the design of neighborhood safety applications [18]. Other neighborhood safety applications have also been explored in the HCI community [42, 79]

Within the broader field of HCI, there is a discernible need for scholarship that engages with the field's growing social justice research agenda. To this end, scholars have explicitly called for an emphasis on exploring the intricate challenges associated with platform-based surveillance and gentrification [19]. Corbett and Loukissas [19] underscore the role of discourse in gentrification and its connection to space, place, and technology. In their work, the authors present three examples of how socio-technical systems mediate consumption-driven gentrification, including Yelp, Nextdoor, and Zillow. While their paper does not establish a direct causal link between these platforms and gentrification, the authors enumerate the ways in which HCI research can engage with such issues [19]. We contribute to this growing corpus through a multi-layered analysis of posting patterns, and how offline gentrification trends are reflected in the types of posts made to Nextdoor.

3 STUDY DESIGN

Our study is grounded in data analysis of posts made to Nextdoor neighborhoods in Atlanta. In this section, we provide context for the site of our research and the data collection process. We detail the steps taken to filter posts for relevance to *community surveillance*, map neighborhood metadata, construct and apply our adapted codebook to tag posts using automated methods, and locate emerging trends for our theoretically informed typology.

3.1 Context

3.1.1 Nextdoor. Headquartered in San Francisco, California, Nextdoor was founded in 2008, and is available in 11 different countries [9]. Signing up for a Nextdoor account is not as simple as installing the application and creating an account with an email. To gain full access to the platform's functionality, each Nextdoor

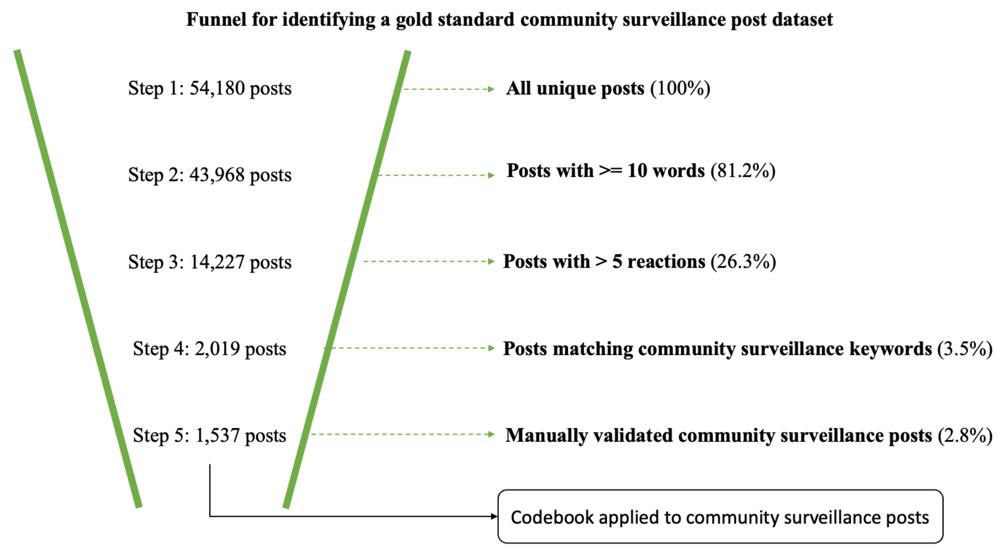


Figure 1: The process for filtering the full Atlanta Nextdoor post corpus to a dataset of 1,537 community surveillance posts.

user is required to authenticate their address via postal mail [62], and the platform encourages the use of real names, rendering the platform a space for de-anonymized users [59]. Users post content (text, photos, videos) on Nextdoor, similar to how posts work on popular social networking sites, and these posts are seen by others in the user's given neighborhood and surrounding neighborhoods. Similar to other popular social networking sites, users can react to posts on Nextdoor, choosing from the standard "like" (depicted by a heart emoji), or indicating "thank you," "agreement", "haha," "wow," or "sad."

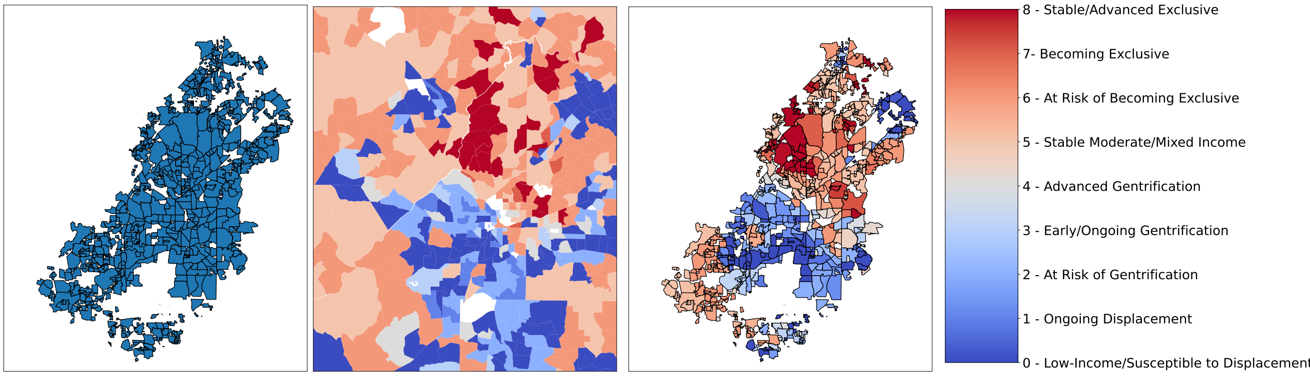
3.1.2 Neighborhood: Atlanta, Georgia. Atlanta is a critical site in the history of race relations in the United States due to its pivotal role in the struggle for racial equality [33, 43]. As such, our rationale for this geographical focus in this study is a convergence of historical significance, theoretical and scholarly precedent, and technical constraints. First, Atlanta has experienced rapid urban development and gentrification in certain neighborhoods, which has led to significant changes in its social fabric and raised issues related to displacement, inequality, and community identity [43]. Multiple prior studies have analyzed local social spaces in Atlanta. As such, existing work on Nextdoor and the social crime-reporting app Citizen conducted in Atlanta provide a backdrop for our research [18, 59]. Second, Atlanta is a demographically diverse city with high income inequality, which prior work suggests may influence the creation and distribution of crime and safety posts on Nextdoor [39, 51]. The city's modern demographic makeup is diverse, with a substantial African American population and a growing number of Hispanic and Asian residents [18, 43]. This diversity has made Atlanta a microcosm of the complex racial and socioeconomic dynamics found in many American cities. Multiple prior studies have analyzed local social spaces in Atlanta. Third, there is detailed geographic neighborhood-level gentrification data available for the city of Atlanta that we can integrate into our analysis [41].

3.2 Data Collection and Filtration

For our study, we gather user-generated content from Nextdoor to compile a corpus of posts made to neighborhoods in Atlanta, Georgia. We gather a collection of posts most likely to be related to community surveillance. We manually assess whether each post should be included in a gold-standard dataset of high-impact community surveillance posts. The selection process is summarized in Figure 1.

To gather posts, we leverage Nextdoor's directory of all neighborhoods to collect posts for all Nextdoor neighborhoods in Atlanta. Our data set of posts date from March 1 2022 to March 1 2023, across 716 neighborhoods in Atlanta. In addition to the post text, we also gather additional metadata such as the date of the post, the number of comments, the number of reactions, the top reactions, and whether the post includes media such as a picture or video. We extract post data from Nextdoor using the python *selenium* package, and use the *beautifulsoup* package to extract data from the HTML. Measures were taken to ensure the speed of data collection would not negatively impact Nextdoor's servers: a 30 second buffer was included between each neighborhood collection, and only one window was open at a time. Data collection took approximately 30 days.

After collecting posts made within a one-year time period across all Nextdoor neighborhoods in Atlanta, our dataset amounted to 54,180 unique posts across 651 Atlanta Nextdoor neighborhoods. As a pre-processing step, we narrow our dataset to only include the 43,968 posts that consisted of at least 10 words. Many posts with fewer than 10 words were media-only posts which included only pictures, videos or links.



(a) Nextdoor neighborhood shapefiles **(b) Urban Displacement Project gentrification map** **(c) final gentrification map for Nextdoor neighborhoods**

Figure 2: Figure depicts the process through which we associated a Nextdoor neighborhood with its corresponding gentrification index. Figure 2a shows the original Nextdoor Atlanta neighborhood shapefile, Figure 2b shows the Urban Displacement Project shapefile, and Figure 2c shows the overlap and resulting final gentrification index score for each Nextdoor neighborhood.

We conduct an initial layer of filtering to narrow our dataset to a subset of likely relevant posts. Guided by our theoretical framing, we consider a community surveillance post as any post that is related to enforcing community norms, surveillance, safety, and crime. The first layer of filtering we perform focuses on posts which received five or more reactions from other Nextdoor users. This choice was motivated since prior work (e.g. [75]) and our own observations of the corpus substantiated the idea that posts related to crime are more likely to receive engagement on social media. We also focus on posts with more reactions as a signal that these posts were the most impactful on the platform. After filtering to posts with more than five reactions, the corpus consisted of 14,227 posts.

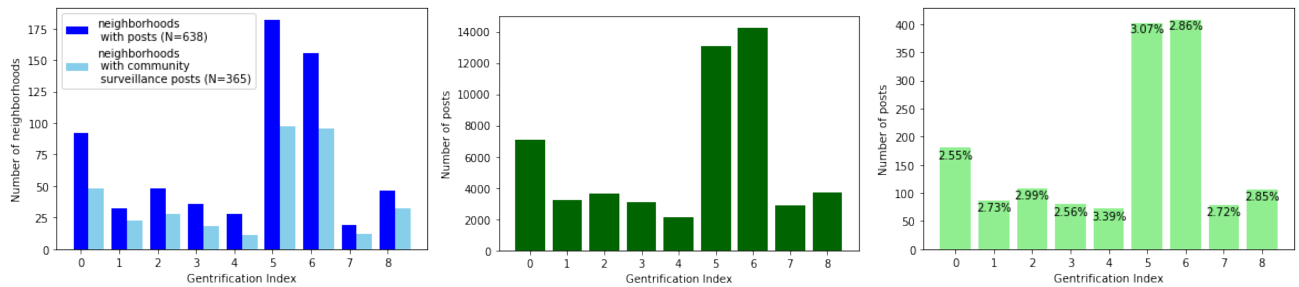
We employed a semantic model to further filter for posts most relevant to community surveillance. Our model of choice, S-BERT, is a sentence embedding model that allows us to encode both a source phrase (the Nextdoor post) and a target phrase (a community surveillance keyword), and measure their similarity. We use the pre-trained *msmarco-distilbert-base-v4* S-BERT model, which is optimized for social media data [70]. We use the cosine similarity between the Nextdoor post and target keyword as our similarity metric. This method has been shown to successfully identify posts related to crime on Nextdoor using a cosine similarity threshold of 0.7 or greater [39]. Since we include an additional manual filtering step, and thus recall is more important than precision for our uses, we settle on a more lenient value of 0.2 or higher.

The S-BERT method relies on a validated set of surveillance-related target keywords. The keywords were arrived at through iterative labeling of posts by three researchers on this project, followed by removal of highly correlated keywords. This process yielded 17 community surveillance-related keywords. A post was considered relevant to the keywords if the cosine similarity between the post text was 0.2 or greater. Using this mechanism, a total of 2,019 posts were retained for further analysis. More information about the keyword selection process and their distribution are included in the Appendix.

As a final step in the filtration process, we create a gold standard dataset of community surveillance posts for further labeling. Leaning on our definition of community surveillance posts as “any post that is related to enforcing community norms, surveillance, safety, and crime,” the two lead authors separately coded a sample of 100 filtered posts. The coders were in perfect agreement over which should be included or excluded as “community surveillance” posts (Cohen’s $K = 1$), thus we could proceed with independently labeling all data. The research team manually labelled 1,537 posts of the 2,019 originally filtered posts (76%) as related to community surveillance. The high final inclusion rate also serves as validation that our selection process was a reasonable proxy for filtering community surveillance posts. We include additional data validation steps taken in the Appendix. Further analyses are conducted using a final dataset of 1,537 community surveillance posts from 365 Nextdoor neighborhoods.

3.3 Mapping Neighborhoods to Metadata

To understand how neighborhood displacement trends may impact hyperlocal social media posts, we associate each Nextdoor neighborhood to its corresponding gentrification index. We build on a geographic gentrification map for the city of Atlanta developed in 2018 by the Urban Displacement Project (UDP), a research action initiative that aims to “understand and describe the nature of gentrification, displacement, and exclusion” [41]. The researchers at UDP developed a typology of nine ordinal categories of neighborhood urban displacement from *low-income/susceptible to displacement* through *stable/advanced exclusive*. We select this data source since it is the most fine-grained available dataset of Atlanta gentrification trends, and was constructed in collaboration with local community organizations. We retrieve the shapefile of each Nextdoor neighborhood, which allows us to identify the geographical area and coordinates. We then match each Nextdoor neighborhood to the geographical index of gentrification in Atlanta from the UDP. Where the shapefiles intersect, we take a weighted average of all indices covered within the Nextdoor neighborhood. This mapping process



(a) # of neighborhoods **(b) # of posts** **(c) # of surveillance posts**

Figure 3: This figure visually demonstrates high-level neighborhood and post characteristics. We observe a similar distribution across all variables, where there are more neighborhoods and posts for neighborhoods with a gentrification index of zero, and for a gentrification index of five or six. Subfigure (c) demonstrates that the % of posts tagged as surveillance posts is reasonably consistent between neighborhoods with different gentrification indices.

is depicted in Figure 2. Since we have relatively limited neighborhoods for some indices, we simplify the 9-step gentrification index to three main categories that we apply to each neighborhood. In our analysis, a Nextdoor neighborhood is considered either "not gentrified" (low-income/susceptible to displacement, ongoing displacement, at risk of gentrification), "gentrifying" (early/ongoing gentrification, advance gentrification, stable moderate/mixed income), or "exclusive" (at risk of becoming exclusive, becoming exclusive, stable/advanced exclusive). Using this mapping procedure, we are able to assign a gentrification index to 703 out of 716 Atlanta neighborhoods we collected. The non-tagged neighborhoods are those which the UDP has assigned a "neutral" gentrification category, such as "High Student Population," or "Unavailable or Unreliable Data."

While RQ2 concerns gentrification, we also associate Nextdoor neighborhoods to additional offline demographics to further contextualize our results. We retrieve each neighborhood's offline demographics using data from the 2021 American Communities Survey (ACS) run by the Census Bureau[1]. To map the Nextdoor neighborhoods to the ACS data, we calculate the percentage of each Zip Code Tabulation Area (ZCTA) using the same weighted average area calculation as described above. We collect information about neighborhood-level population, sex, age, income, race, and education. We note that the total number of ZCTAs in the Atlanta area is lower than the more fine-grained UDP gentrification data, with 61 ZCTAs overlapping with our dataset compared with 264 neighborhoods in the UDP.

As a first step, we visualize the distribution of Nextdoor neighborhoods relative to their gentrification index in Figure 3 to descriptively explore the Nextdoor neighborhoods in Atlanta, GA. The subfigures illustrate that there are comparatively more Nextdoor neighborhoods in Atlanta, GA that are classified as "Stable/Moderate Income" and "At Risk of Becoming Exclusive." Figure 3b demonstrates that the number of posts is reasonably proportional to the number of neighborhoods, and Figure 3c shows that the percent of community surveillance posts across gentrification indices is comparable ($min = 2.56\%, max = 3.39\%$). In all, there are between 72 to 408 posts made to neighborhoods of each gentrification index in the final dataset.

We check for potential issues of bias in our data in the selection of posts and in the resulting representation of neighborhoods. Figure 3c suggests that the proportion of surveillance posts is relatively constant, and the number of posts appears to reflect the overall distribution of neighborhoods across the gentrification index. We used statistical tests to confirm this intuition. We run a two-tailed t-test to compare between the distributions of the gentrification index of neighborhoods that have been selected for final analysis ($M = 4.32, SD = 2.43$) and neighborhoods that have not ($M = 4.07, SD = 2.34$), and find no significant difference between the two groups, $t(701) = 1.4, p = 0.17$. We also run a two-tailed t-test to compare between the distributions of the gentrification index of posts that have been selected for final analysis ($M = 4.36, SD = 2.35$) and posts that have not ($M = 4.31, SD = 2.40$), and find no significant difference between the two groups, $t(1633.2) = 0.72, p = 0.47$. We note that these results are congruous with the belief that there is bias in discussions of crime on Nextdoor based on income – if crime discussions on the platform were exactly reflecting offline crime rates, we might expect to find that poorer neighborhoods have higher rates of community surveillance posts. Like in prior work, we find that the number of posts in a neighborhood strongly correlates with the local population, $r(701) = 0.53, p < 0.0001$. [39]. In sum, we do anticipate that areas with higher populations will be proportionally more represented in our dataset, but we do not anticipate that our selection method has introduced further bias relative to neighborhood gentrification.

3.4 Codebook Development

Our methodology combines both deductive and inductive approaches to qualitative coding. We develop and iterate on a codebook, then evaluate GPT-4's [64] performance against the codebook.

In this work, we demonstrate the use of GPT-4 for large scale qualitative coding. We explored using GPT-4 for this analysis since we sought to develop an extensible framework that demonstrates using automated methods to code Nextdoor (and other) text, and

[1] https://www.census.gov/data/2021/

large language models are an emerging area for tagging natural language text [6, 31, 76]. One way GPT-4 is preferable to crowdworkers is preventing trauma and mental harms which can result from exposure to posts about topics such as crime and violence [16, 21]. Enabling the automatic labeling of challenging topics can limit the trauma put upon data workers, which may be used for future research in important societal topics such as discrimination, violence, or mental health. We leverage GPT-4 for coding and showcase its successful application as an emerging approach in mixed method research. We develop this approach to flexibly enable future work and larger-scale analysis, most promisingly going beyond one location to perform this analysis in other cities. In the future, this framework can be used by other researchers conducting HCI research on social platforms.

We first develop a codebook for human labeling, which we then adapt to function for GPT-4. To develop our codebook, we took inspiration from López and Butler's codebook developed for tagging posts made to local Facebook groups, which includes tagging what a post is about, labeling the intent of the message, and inferring what is requested from the audience [56]. In addition, three researchers inductively assessed a set of 25 community-surveillance related posts to identify emerging themes. We combine these into a first version of the codebook that captures the poster's stated identity, the subject of a post, the intent of the message, the primary sentiment, and key governance tactics employed. In the first round of coding, we manually labeled a subset of 340 filtered community surveillance posts.

We then tested for the feasibility of leveraging LLMs to label the posts. We develop an adapted codebook prompt to provide to GPT-4 and to use as researchers when performing the coding. Many changes were made between human and GPT-4 tagging instructions. The GPT-4 codebook required more precise language. For example, the human raters easily agreed on what should be deemed "explicitly calling for vigilance," whereas the GPT-4 codebook had to be rephrased to the more exact "explicitly asking people to be cautious or alert." Three members of the research team then coded 100 posts each: 50 posts were overlapping between all three researchers, 50 posts differed. Using crude agreement, we confirmed that GPT-4 coded approximately as well as the three human coders and refined the codebook one final time.

In the third round of coding, two researchers reached high interrater agreement. GPT-4 achieved a Cohen's κ between 0.65-0.81 for all but two codes, which were deemed sufficient for further analysis. The final codebook and inter-rater agreements between all raters are attached in the Appendix. To enable future research, we attach the prompt passed to GPT-4 in the supplementary materials.

3.5 Typology Formation

To form the typology of Nextdoor community surveillance posts, we take a mixed methods approach. We draw on our qualitative observations of the data while coding, and use statistical methods to visualize common relationships between posts.

To visualize relationships between the post labels, we employ a method called Multiple Correspondence Analysis (MCA). MCA allows us to reduce our qualitative codes to fewer dimensions that capture a large portion of the variation within our codes [1]. MCA can be used as a method to explore relationships between quantitative and qualitative data in the same multi-dimensional space, and to group observed categorical variables together [20, 22, 28, 46]. The unit of analysis for the MCA is at the post-level, meaning that the MCA is constructed from the 1,537 community surveillance posts and associated metadata. An MCA consists of active variables, which are directly used to build up dimensions, and supplementary variables, which can be used to describe the resulting dimensions. As the active variables to construct the MCA dimensions, we use the codebook tags assigned by GPT-4 to each post to generate our dimensions for analysis. To understand whether neighborhood gentrification trends correlate with the model dimensions, we include the simplified neighborhood gentrification index as a supplementary qualitative variable, meaning each post is labelled as being posted to a neighborhood that is either "not gentrified," "gentrifying," or " exclusive." We also include other descriptive data as supplementary variables to further contextualize our results, including the census data and platform metadata such as the number of comments.

After interpretation of the MCA results and discussion of our codebook, we converge on a typology of community surveillance posts. We identify three distinct clusters of posts in the first two dimensions of the MCA, which became the basis for the three larger categories of posts we label. To interpret patterns of community surveillance within each cluster, we observe where the qualitative and quantitative codes are situated in the two-dimensional space. We also reference the co-occurrence matrices between different codes included in the Appendix. Throughout this constructive process, we also examine and reference individual posts in the two-dimensional space.

4 POSITIONALITY AND ETHICS

Ethical considerations have been prioritized throughout this project. In this section, we provide a positionality statement where we consider the ethical implications of our methodology, and summarize privacy measures taken to minimize harms.

4.1 Author Positionality

We acknowledge our own positionality relative to this study. The authors involved with this work represent voices from multiple countries, genders, races, and varied socioeconomic circumstances. There are two primary female authors for this paper whose perspectives are most present in this paper. Madiha is a first-generation ethnic minority who is an expert on community norms of information sharing in online spaces. Marianne is an expert on local communities, and has investigated other hyperlocal platforms such as Facebook groups, local subreddit communities, and local news media. The authors have complementary expertise and this research is a collaboration between a qualitative and quantitative perspectives.

While we are not experts in the specific geographic area this study concerns, we selected this space for both contextual and practical reasons outlined in the methods and guided by prior works. We acknowledge that this study has implications specifically for black/white social divides, and specific implications for the area of Atlanta. Our approach towards this work is to center and respect

Table 1: Example of the post scrubbing process on mock data.

Example post before scrubbing
Hello, my name is **Annie**, I just moved to the neighborhood on **8**th Ave and Orchard Street. My email is **annie@gmail.com**, I can't wait to meet everyone at my housewarming party on August **8**th at **7**pm!

Example post after scrubbing
Hello, my name is **#PERSON#**, I just moved to the neighborhood on **#NUMBER#**th Ave and Orchard Street. My email is **#EMAIL#** I can't wait to meet everyone at my housewarming party on August **#NUMBER#**th at **#NUMBER#**pm!

the experiences of both the people who are posting on Nextdoor, and those who are being posted about without consent.

4.2 Collecting Nextdoor Data

We do not take lightly that collecting social platform data reifies the systems that produce toxic content and surveillance capitalism. While journalists hold that accountability through methods such as web scraping are "vital for democracy," [73] quantitative data aggregation can easily become a tool that reinforces existing systems of oppression [24]. In addressing our choice to collect and utilize Nextdoor data, we underscore the timely necessity of our unique qualitative findings within the broader landscape of social platform research. Journalistic evidence documents the presence of community surveillance phenomena on Nextdoor. However, these narratives concern highly specific events across a wide variety of geographic areas that focus on platform toxicity and discrimination [37, 48, 52, 78]. We treat media investigations as motivation, but underscore that they do not establish a process for tracing how community surveillance occurs at the post level, what norms emerge from these patterns, and the relationship between community surveillance and gentrification trends as our work seeks to do. Our study connects existing quantitative analyses of Nextdoor content with theoretical qualitative discourse on surveillance and gentrification. We uncover the specific patterns and trends of community surveillance that emerge within online local communities. Further, our typology of community surveillance posts serves as an evidentiary and validated framework that is extensible to a broader spectrum of online platforms. We believe this work can be influential to the CHI community beyond implications for the Nextdoor platform.

4.3 Preserving User Privacy

Due to the verification process required for a Nextdoor account, Nextdoor users likely have an expectation of relative privacy within their neighborhood. In the absence of publicly available Nextdoor data for research, we take a number of steps to protect user data and anonymity. We believe our safeguards help to obscure any identifiable information that researchers may have gathered from Nextdoor users.

First, we only gather public posts that are marked with the visibility set to "anyone." Since "anyone" is the default setting on Nextdoor posts, we believe a majority of posts on Nextdoor are likely to fit under this description. However, users may choose to tag their post as visible to "your neighborhood" or "nearby neighborhoods" only. We never gain access to an account holder's full name. Nextdoor makes users' first and last initial available to users who have not validated their location. Only a hash of these initials were stored, such that we cannot decipher the initials of the poster. We use regex and natural entity recognition to match and scrub potential phone numbers, other numbers, email addresses, and names prior to saving any of the posts. Table 1 shows a mock example of this process. We also paraphrase all posts included in this external analysis, following recommended best practices [3].

We take additional precautions before processing data using a large language model, given the emerging concerns around conducting research with proprietary models [63]. Numbers, phone numbers, email addresses, and names had already been removed from the posts. We judged that the primary remaining threat to participants was revealing specific locations, for example by specifying an intersection of two streets. The research team manually combed through all relevant posts to judge whether they might benefit from additional anonymizing. Among the set of 2,019 posts, 226 (11.2%) were further anonymized. When unsure, we defaulted to anonymizing.

We do not believe any of the users of Nextdoor to be at risk of identification or harmful impacts as a result of this work. Our study meets the Safe Harbor Rules for Collected Data §164.514(A) of the HIPAA Privacy Rules). This protocol was also reviewed and marked exempt by Anonymous Institution's IRB. Our privacy strategy strives to model ethical data collection for the broader HCI community whose work seeks to expose structural inequalities perpetuated by platforms.

5 FINDINGS

In our findings, we first summarize the codes assigned to the community surveillance posts, followed by visualizing all codes in two dimensions using the MCA method. We synthesize the findings into a typology of community surveillance patterns uncovered by user-generated posts on Nextdoor.

5.1 Summarizing community surveillance posts

We summarize the categories assigned by GPT-4 to the 1,537 identified community surveillance posts on Nextdoor in Table 2. From this analysis, we determine that the largest categories of Nextdoor community surveillance posts are *criminal or suspicious person* (19.8%), *not applicable* (16.4%), *theft* (15.4%), and *inanimate object or animal* (13.1%). The majority of community surveillance posts describe a user's personal experience, use a negative sentiment, and approximately a third of all posts explicitly call for vigilance. Approximately one-tenth of all posts include a physical description of one or multiple people.

5.2 Visualizing community surveillance posts

The multiple correspondence analysis (MCA) technique allows us to visualize the qualitative codes from our codebook in multidimensional space, such that codes that resemble one another occupy a similar area in the X-Y dimensions. For simplicity, we focus

Table 2: Tagged attributes of the 1,537 community surveillance posts.

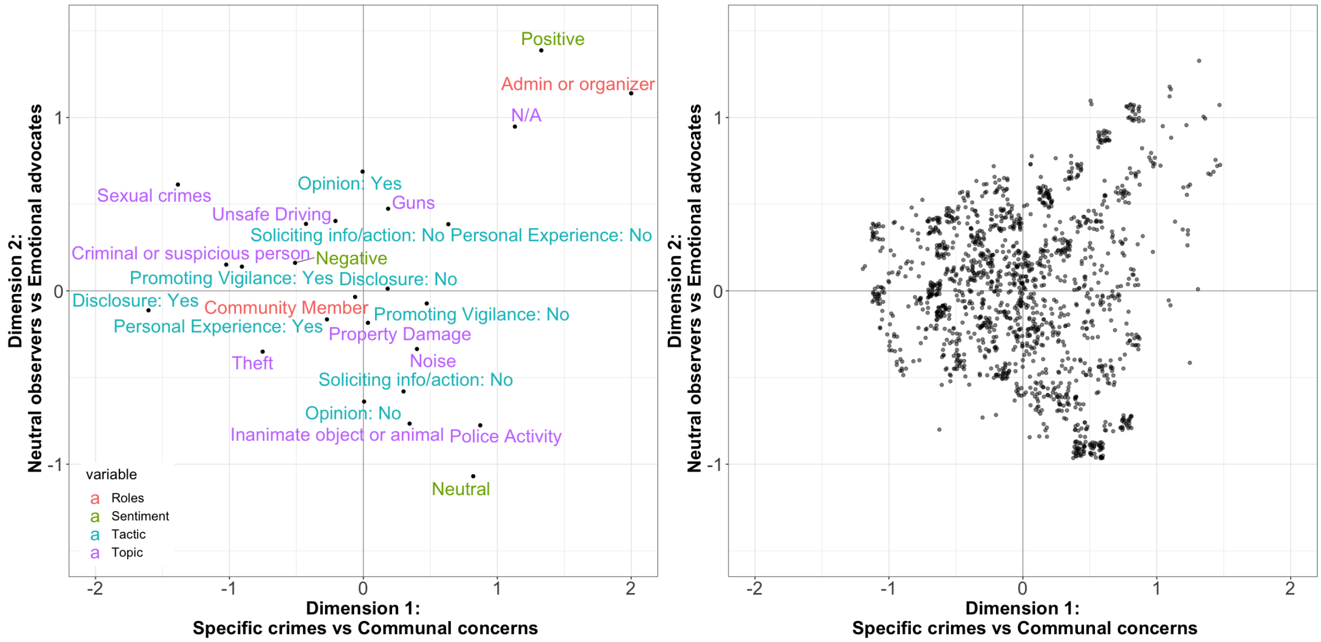
Category	Code	Posts tagged	% Share
Main post topic The central matter of interest in a post	criminal or suspicious person	304	19.8
	not applicable	252	16.4
	theft	238	15.4
	inanimate object or animal	202	13.1
	police activity	170	11.1
	guns or gunshots	136	8.8
	unsafe driving	114	7.4
	property damage	59	3.8
	noise	53	3.4
	sexual violence or harassment	9	0.6
Roles A function assumed by a person posting on the Nextdoor platform	community member	1491	97.0
	administrator, organizer, or moderator	46	3.0
Describing personal experience Anyone who is sharing an event they witnessed first-hand or that happened to them	yes	1076	70.0
	no	461	30.0
Expressing personal opinion Someone sharing their viewpoint on something without being prompted or with the intent to convince others	yes	740	48.1
	no	797	51.9
Soliciting information or action Explicit requests for something, e.g. pictures of an incident	yes	630	41.0
	no	907	59.0
Calling for vigilance A post that explicitly asks people to be cautious, watchful, or be on the lookout for someone.	yes	527	34.3
	no	1010	65.7
Physical description The post describes how a person looks, e.g. their race, age, or gender.	yes	156	10.1
	no	1381	89.9
Primary sentiment The main emotion a post is likely to evoke in the reader	positive	163	10.6
	neutral	364	23.7
	negative	1010	65.6

this analysis on the first two dimensions, though we also explored other dimensions to help us construct our typology. The model had 17 dimensions overall, and dimensions 1 and 2 together explained about 23% of the variance. At a high level, Dimension 1 appears to distinguish between posts about specific crimes or incidents committed by one or more individuals (*specific crimes* ⟵), and posts targeted at the general neighborhood (*communal concerns* ⟶). Dimension 2 appears to separate posts that are made from a positive or negative perspective (*emotional advocates* ↑) from posts that are making neutral or curiosity-based statements (*neutral observers* ↓). The first and second dimension are visually displayed in Figure 4.

Dimension 1, *specific crimes* ⟵ *vs communal concerns* ⟶, had an eigenvalue of 0.279 and described 13.1% of the variance. The codebook variables *main post topic*, *primary sentiment*, and *explicitly promoting vigilance* best describe this dimension. Figures 5a, 5b and 5c show how these codes divide dimension 1. Posts closer to the *specific crimes* ⟵ side tend to have a negative sentiment, explicitly promote vigilance, and publicly disclose a physical description. In this dimension, posts that are closer to *communal concerns* tend to not be explicitly promoting vigilance, not be disclosing public descriptions, and the main post topic is not identified.

Dimension 2, *emotional advocates* ↑ *vs neutral observers* ↓, had an eigenvalue of 0.204 and described 9.6% of the variance. The codebook variables *primary sentiment*, and *expressing personal opinion*, and *main post topic* best describe this dimension. Figures 5a and 5d show how these codes divide dimension 2. In this dimension, posts that are closer to *emotional advocates* ↑ tend to be expressing a personal opinion, do not tend to be soliciting an information or action, and have a positive sentiment. Posts closer to the *neutral observers* ↓ side tend to not be expressing a personal opinion, have a neutral sentiment, and do solicit information or action.

Another goal of our analysis was to explore how gentrification may impact the nuanced types of community surveillance posts made by users in certain neighborhoods. To this end, we report the



(a) Visual representation of codebook **(b) Visual representation of individual posts**

Figure 4: This figure visually demonstrates the results of the MCA along the two dimensions. Dimension 1, from left to right, appears to separate posts about *specific crimes* ← from *communal concerns* →. Dimension 2, from top to bottom, appears to distinguish posters who position themselves as *emotional advocates* ↑ vs. *neutral observers* ↓. Figure 4 a shows how the codes relate to each other in this two-dimensional spaces. Figure 4b shows where the individual posts lie, and appear to show three clusters of posts: one that primarily occupies the left of the axis, and two that are in the top and bottom right quadrants.

(a) Primary sentiment **(b) Calling for vigilance** **(c) Physical description** **(d) Personal opinion**

Figure 5: This figure visually demonstrates how the data is split among four of the most representative codes for dimensions 1 and 2. Figures 5b and 5c demonstrates that calling for vigilance and giving a physical description are associated with *specific crimes* ←. Figure 5d shows that giving a personal opinion is associated with posts made by *emotional advocates* ↑.

correlations between the gentrification index of the neighborhood the post was made in and dimensions 1 and 2 in Table 3. Neighborhoods that are "not gentrified," and likely to be lower income, are significantly positively associated with dimension 2, meaning they correlate with *emotional advocates* ↑ posts on Nextdoor. Neighborhoods that are "gentrifying" are significantly negatively associated with dimension 1 and negatively associated with dimension 2, meaning they correlate with *specific crimes* ← and *neutral observers* ↓ posts on Nextdoor. Neighborhoods that are "exclusive" are only significantly positively associated with dimension 1, meaning they are correlated with *communal concerns* →. Notably, as the gentrification index increases, the correlation vector moves from pointing upwards, and slightly left, to the lower left, to the right in a counterclockwise motion; indicating the community surveillance posts may shift from top left to the right in ways that reflect their underlying socio-economic environment. The vectors are further visually

Table 3: Table shows how neighborhood gentrification may relate to the two dimensions in the MCA. Overall, neighborhoods that are "not gentrified" correlate with the *emotional advocates* ↑ direction, "gentrifying" neighborhoods correlate with *specific crimes* ← and *neutral observers* ↓, while "exclusive" neighborhoods correlate with *communal concerns* →.

Neighborhood categorization	Dim 1	Dim 2	Vector direction	Correlated with
Not gentrified	-0.03	0.12^{**}	↑	*emotional advocates*
Gentrifying	-0.17^{**}	-0.18^{***}	↙	*specific crimes* *neutral observers*
Exclusive	0.09^{***}	0.01	→	*communal concerns*

Note: *p<0.05; **p<0.01; ***p<0.001

depicted in the Appendix. Though these correlations should not be taken as definitive proof that certain types of posts are only made in one type of neighborhood, they do help to assess trends in how community surveillance posts on Nextdoor may morph as a neighborhood becomes more gentrified.

Although the primary focus of this study and our research questions concern gentrification, we add further context to the analysis through mapping additional demographic variables, platform-level variables, and the community surveillance keywords onto the MCA's two-dimensional space. While Census data is informative, the data is less fine-grained than the gentrification data and thus we caution against taking these findings as definitive for such a localized analysis. Notably, the census variables bisect the data in dimension 2 more than in dimension 1. A higher proportion of Black or African American residents and a higher poverty percentage in a neighborhood is significantly correlated with the *emotional advocates* ↑ direction, and a higher percentage of men, higher population, higher median income, more educated population, and whiter population is significantly correlated with the *neutral observers* ↓ direction. Finally, the number of reactions made to a post is significantly associated with the *emotional advocates* ↑ dimension while the number of comments is associated with both the *emotional advocates* ↑ and the *specific crimes* → direction. In terms of the reactions, the *like* reaction is significantly associated with the *communal concerns* → direction, while the *wow* reaction is significantly associated with the *specific crimes* → direction, and the *sad* reaction is significantly associated with the *specific crimes* ← and *neutral observers* ↓ directions. More information about how our dimensions map on to census data is included in the Appendix.

We also checked for potential biases in the data. We modeled the percentage of community surveillance posts out of all posts in a neighborhood as a quantitative variable, and find that this variable is *not* significantly associated with dimension 1 or dimension 2 of the MCA. This finding suggests that the dimensions are not significantly influenced by whether a neighborhood contained more community surveillance posts.

5.3 Typology of Community Surveillance Posts

Using the qualitative and quantitative analysis methods described, we present a typology of community surveillance posts on Nextdoor to organize the emerging patterns we uncover. A majority of community surveillance posts on Nextdoor are 1) incident reports, 2) community norm formation posts, or 3) informative posts. Although these categories roughly map to the three clusters identified from the MCA, they are more nuanced and not a direct interpretation of the dimensions. We discuss the distinguishing themes present within each of these posts below. We include examples from our data, which are manually paraphrased so that they cannot be searched, details have been removed, and in some cases represent amalgamations of multiple similar posts. An overview of these types is presented in Table 4.

Table 4: Typology of community surveillance posts on hyperlocal platforms

Post Type	Theme
Incident report	Sharing personal experience Calling for vigilance Public disclosure Crowdsourcing action Witnessed incident Information gathering
Community norm formation posts	Correcting behavior Demanding action Reinforcing positive norms
Informative posts	Sharing local news reports Authority figure or organizer Knowledge construction

5.3.1 Incident reports. In this type of community surveillance post, the poster describes a criminal or normatively transgressive incident involving one or more specific offenders that the poster has usually witnessed or experienced. The themes relating to this post type are *sharing personal experience*, *calling for vigilance*, *public disclosure*, *crowdsourcing action*, *witnessed incident*, and *information gathering*.

The poster is frequently, but not always, the victim of an incident. Sometimes, the poster simply shares their experience, seemingly looking for sympathy, reassurance, or feeling unsure what other action to take. These types of posts seem to occur often when posters are feeling some level of impotence – such as when the poster is not sure an incident warrants police involvement, or when reporting their experience to law enforcement feels insufficient.

> Sharing personal experience: *My daughter and I experienced a frightening incident today when some boys were throwing rocks at us near [place]. The rocks hit our car and damaged the windshield. we're both okay but it was a scary experience, the rocks narrowly missed my daughter..I'm grateful to God it wasn't more serious.*

Often, the poster calls for a state of alertness from their neighbors, sometimes in conjunction with a physical description or media of the suspect(s). The posters seem to believe these types of posts will help protect the community against such crimes or activity.

> Calling for vigilance: *Kid attempted to break into cars 15 minutes ago. Police notified. Please be on the lookout.*

A related theme within these posts is the tendency to publicly disclose information about people deemed suspect. While only 10% of the community surveillance posts we analyzed included this tactic, about 19.6% of the posts considered in the *specific crimes* ← pole of dimension 1 included some sort physical description. Notably, these physical descriptors are often not enough for a reader to uniquely identify the people being described.

> Public disclosure: *there are these two guys hanging out at the back of [street]. asked them to leave, but they didn't have a ride. They tried bothering an elderly neighbor, but she's smart and didn't answer the door. they're still here. one's wearing a navy shirt, the other's a dark-skinned guy with a beard, wearing a blue shirt.*

Other times, the poster requests specific action from the community, such as calling the police, sharing any security footage of an incident, or getting in touch if readers have pertinent information.

> Crowdsourcing action: *Just a heads up, this [incident] happened around [place] drive. If you happen to have any cameras in that area, please hit me up. And if you've seen a guy wearing this sweatshirt around there, please let me know. [image of sweatshirt]*

In some cases, the poster states they have information about an incident that did not affect them, but may impact someone in the community. These posts are made in case the information is helpful to someone else or in an effort to reach those impacted people.

> Witnessed incident: *I saw someone hit the mailbox on [ave] just after [street]. If you want more details, please send me a DM.*

Finally, posters sometimes look for further information about a known incident they have little relationship with. In these cases, the poster has heard of or partially witnessed an event in the area, but is looking for more details.

> Information gathering: *I'm wondering if anyone has any information about the girls who were found dead in a hotel in the [place] area.*

5.3.2 Community norm formation posts. In the second type of community surveillance post, the poster perceives a trend of behaviors occurring in the neighborhood that they feel may be relevant to community members, which they wish to reinforce or alter. The themes associated with this post type are *correcting behavior*, *demanding action*, and *reinforcing positive norms*. These posts exhibit a number of motivations, but are often associated with stronger emotions from the poster. The first theme in this type of community surveillance post emerges when the poster wishes to correct neighborhood behavior, conditions, or a trend in their community which they find unacceptable.

> Correcting behavior: *It's really getting on my nerves how some people keep revving their engines so loudly in the tiny parking lot of [business]. You can hear it all the way up to the apartments and it's seriously disturbing the peace.*

The poster may also demand action if they have observed behaviors, conditions, or a trend in their community which they feel are unacceptable. In these posts, the user explicitly recommends a suggested policy or action from government, companies, or individuals. These posts can be emotional, and suggest pent-up frustration at the situation.

> Demanding action: *There is more evidence that we need speed bumps on [intersection]. Too many people drive dangerously through the stop signs. Today someone sped through the stop sign and hit the curb, hitting a tree, tore up their car and the airbags went off.*

Some of these types of posts also attempt to reinforce positive norms. For example, if a poster has experienced or observed a positive action in the neighborhood, they express their gratitude online on the Nextdoor platform, and make a positive example of their story. This interaction reinforces a positive norm of gratification via oversharing. In our dataset, only about 10.6% of posts are overall expressing a positive sentiment.

> Reinforcing positive norms: *Big thanks to Officer [name] and [name] from the [place] Police Department for coming to our rescue when we ran out of gas on the way to the gas station. They were real lifesavers!*

5.3.3 Informative posts. The third type of community surveillance posts we identify are informative posts, where Nextdoor users possess information they view as important to pass along to the broader community for seemingly altruistic motivations. The themes associated with this post type are *sharing local news reports*, *authority figure or organizer*, and *knowledge construction*. As a way to pass along information, posters share or paste a link to a local news report, usually without their own commentary. In our dataset, these news reports are most often about crimes, police work, or local authority figures.

> Sharing local news reports: *Teenager found dead following car shooting outside Atlanta police station*

Another theme among informative posts are posters who identify themselves as organizers or community figureheads, for example heads of a local business, journalists, or a non-profit that is passing along information to the community. These may also be calls for participation in local community events, for example gun safety or neighborhood watch programs.

> Authority figure or organizer: *We fell short of our high standards in the latest inspection. We fixed issues, retrained staff, and will take extra precautions to prevent recurrence. Sorry for any inconvenience and thank you for your support.*

Lastly, sometimes posters do not present themselves as authority figures, and may not feel that they have all the information, but want to contribute newly acquired information to the conversation so that the community together can arrive at knowledge.

Knowledge construction: *There's been a porch thief on our street. I've seen some of the other posts – maybe the same guy?*

6 DISCUSSION

In this section, we discuss the implications of our findings on communities, and their relationship to neighborhood gentrification. We describe three overarching patterns of community surveillance posts on Nextdoor: *incident reports*, *community norm formation posts*, and *informative posts*. Our analysis suggests that posts made to "gentrifying" neighborhoods on Nextdoor are commonly *incident reports*. Posts made to "exclusive" neighborhoods are commonly *community norm formation* or *informative posts*. Community surveillance posts made to "not gentrified" neighborhoods tend to be more emotional. Such posts either express a personal opinion, touch on violent crime, or relate to both specific crimes and communal concerns.

One contribution of this work is the extensible typology of hyperlocal surveillance posts. While prior work has theorized that gentrification can be furthered by digital hyperlocal spaces [8, 47, 49, 65], we ground our analysis on a large dataset of user-generated content to identify the different *types* of hyperlocal surveillance and how users may employ different tactics to surveil outsiders or promote community behaviors. As an increasing number of works in HCI and related communities wrestle with how to make sense of content shared to hyperlocal spaces [18, 59, 79], this typology helps us reason about such information. Future work may seek to identify posts according to this typology in other hyperlocal or semi-private contexts across platforms.

Our findings extend observed relationships between community income, inequality and how those communities discuss crime and safety [47, 49]. Communal concerns appear to be more closely associated with "exclusive" neighborhoods, while specific crimes are more closely associated with "gentrifying" neighborhoods. Among "not gentrified" neighborhoods, both specific and communal concerns are discussed. Poorer neighborhoods also tend to have higher rates of crime, particularly violent crime [60]. A simple explanation for our finding is that individual crimes are discussed more within neighborhoods where they occur more frequently, the "not gentrified" and "gentrifying" neighborhoods in our data. However, Iqbal et al. found that wealthier Nextdoor neighborhoods discuss crime proportionally more when compared to official crime statistics [38]. Our results, therefore, point to a more nuanced notion that specific crimes continue to be discussed until a neighborhood reaches an "exclusive" status, at which point discussions related to crime become communal in nature.

We also demonstrate nuanced relationships between expressed sentiment and gentrification. Our findings suggest that the sentiment associated with community surveillance posts on Nextdoor shifts from negative to positive as neighborhoods become "exclusive." Using sentiment analysis, Iqbal et al. similarly demonstrated that posts tend to be more positive in wealthier neighborhoods on Nextdoor. One place where our analysis is differentiated is among "not gentrified" neighborhoods. These neighborhoods are more strongly associated in the data with both positive and negative posts than with neutral posts. This finding is corroborated by the census data, where more educated, higher income, and whiter populations are associated with posts tagged as neutral. Speaking about data visualizations, D'Ignazio and Klein [24] make the point that projected neutrality and objective is often "the perspective [. . .] of the dominant, default group." [24]. Specifically, the idea of the white observer in black spaces, who adopts the neutral language of science, has been identified and discussed in prior work [50]. Our analysis may demonstrate this phenomenon again, where the richer, dominant group that perceives itself as objective uses emotionless language to express perceived truth.

6.1 Posting as Policing

The dimensions identified in this study have implications for the ways that online social media are reinforcing socioeconomic gentrification patterns and continuing to exclude minoritized populations. The most common pattern of norm-setting behavior we observe on Nextdoor is the *incident report*. Such reports involve a person sharing their experience with an incident caused by one or more individuals. These posts are frequently accompanied by public disclosure of the suspect's physical appearance (approximately one in five *specific crimes* ← posts). Further, these posts are implicitly applying surveillant and carceral logics, echoing the race- and class-based exclusions historically associated with gated communities [47] where blackness is subject to technological policing [10].

On the Nextdoor platform, *transgressing* in a neighborhood (according to the community's norms) results in an invasion of the *offender's* privacy. In our analysis, we show that posts in "gentrifying" neighborhoods are correlated with the *specific crimes* ← dimension, and posts in "exclusive" neighborhoods are correlated with the *communal concerns* →. Kennedy and Coelho [44] similarly find that members of Nextdoor appear to find outsiders "deserving" of surveillance and monitoring. Once community members have constructed an "in-group", they are capable of othering an "out-group", a process similarly described by Tajfel et al. [74]. Our analysis also shows that in poorer neighborhoods, community surveillance topics that people are more likely to post about concern violent crimes, such as "sexual violence or harassment" and "guns or gunshots," compared with "theft," or "inanimate object or animal." This finding signals that as neighborhoods gentrify, transgressors continue to be punished and othered, but for increasingly petty causes.

Another theme across *incident reports* centers the notion of "boundaries" between a poster, and those deemed by the poster to be a transgressor [25]. *Incident reports* frequently include explicit calls for community vigilance, whereby a poster asks the community to "be careful" or "look out" for potential dangers. These calls echo Foucault's notions of the ominous and ever-present fear of panoptic vigilance [29]. Transgressors are perceived to be outside of the Nextdoor platform, and the Nextdoor community itself, the latter of which posters perceive to be aligned with their interests, and "on their side." Within that frame, the posters are guards, "transgressors" are prisoners and the neighborhood becomes a Foucaultian prison. Posters underscore their victimization. Users perceive the platform as an ally in their victim-hood, rather than a neutral space where community matters are discussed. This process is integral to Nextdoor's design and marketing, not dissimilar to Citizen [18].

6.2 Norms in Exclusive Neighborhoods

Neighborhoods considered "exclusive" are more likely to produce *informative posts* and *community norm-formation posts*. Our findings demonstrate that community surveillance posts originating from "exclusive" neighborhoods display distinct patterns of information sharing and reflect norms specific to these locales.

In "exclusive" neighborhoods, community surveillance posts trend towards discussing communal concerns (such as noise, police activity, or neighborhood hazards). Residents in "exclusive" neighborhoods trend towards reinforcing positive norms of information sharing such as gratitude, requesting more information, or modifying or updating existing posts based on new information. This inclination may be driven by a communal goal to maintain the exclusivity and desirability of their neighborhood, as well as stronger notions of trust among neighbors. This observation points to a potential "hierarchy of needs" of community surveillance, whereby some types of offences must be eradicated before neighborhoods can focus on other types of norm formation. In other words, there is a relationship between gentrification and normatively agreed upon "hierarchy of needs" within a community. When a neighborhood is "not gentrified," instances of both specific crimes and communal concerns are prioritized. In "gentrifying neighborhoods", the emphasis is placed on highlighting specific crimes as they create an obstacle that prevents these neighborhoods from achieving an elevated "exclusive" status. Within "exclusive" neighborhoods, communal concerns become more central. An "exclusive" neighborhood, therefore, can focus more on communal concerns as they no longer have to worry about specific instances of petty crime. Future work may seek to prove out this "hierarchy of needs."

We also find that posts about *communal concerns* → are more likely to solicit information or action from other community members. For example, community members may be unhappy at the state of a local road and solicit others' experiences so they may collectively advocate for change. Prior work has found that hyperlocal social media can be useful to connect members of communities to resources and information, for example during the COVID-19 crises when social platforms could update faster than traditional local news agencies [4, 54, 77]. Calls to exchange information or action recall Putnam's theory of social capital, a term often used as a catch-all for intra-community bonds and trust that lead to mutual benefit [69]. High social capital is more common in higher income neighborhoods [35] and our results demonstrate that these tendencies are also true for communities on Nextdoor. Since posts soliciting information or action are more highly correlated with "gentrifying" and "exclusive" neighborhoods, our empirical evidence shows that Nextdoor can aid the positive construction of social capital – but mostly for wealthier neighborhoods. In a 2001 exploration of internet-based community networks, Jankowski et al. [40] observe that "those geographic communities already rich in social capital may become richer, thanks to community networks, and those communities poor in social capital may remain poor." Our study demonstrates the linguistic strategies that are used to perpetuate the enrichment of already affluent communities.

Community surveillance posts police neighborhoods through a spectrum of tactics. On the punitive end, they can expose the privacy of specific individuals believed to have transgressed, where the offender is perceived as an outsider. On the affirmative side, neighbors can also uphold positive community norms or rally support against neighborhood injustices. Prior work has highlighted that community-based surveillance can just as easily give rise to civilians who are "suspicious" as "civic-minded" [68]. Along these lines, our study empirically demonstrates how this dichotomy plays out on Nextdoor.

6.3 Practical Implications

The first practical implication of our work is related to how Nextdoor may mitigate harms within communities. Our findings and analysis demonstrate communities observe and respond to their neighbors' social cues and that these shifts relate to neighborhood gentrification trends. Nextdoor may be able to regulate the prevalence of divisive content by introducing stronger moderation tools. As described earlier, Nextdoor's "kindness reminder" tool was one attempt to mediate negative sentiments, though largely unsuccessful [37]. To combat "othering" and social exclusion, we see the potential for community-appointed neighborhood moderators representative of neighborhood diversity as one explicit path forward. Recent work in HCI, for example, has started to explore how community values could algorithmically shape feeds [34]. Other moderation tactics such as limiting reactions to crime and safety posts, or explicit reminders or positive affirmations from neighborhood moderators may also more positively guide social cues. Additionally, our research has found, many times over, that the identity of others is frequently disclosed without consent on Nextdoor. The platform could take a stronger stance on detecting and removing the posting of specific names or identity markers from their platform. Ultimately, it is worth acknowledging that not all harms can be prevented: prior work has demonstrated that hyperlocal community-generated content generally is often lower-quality, and reflects the offline structural biases of local communities [5, 55]. In service of understanding where issues do arise, Nextdoor could seek privacy-preserving ways to make research data more available.

Methodologically, we demonstrate the successful use of a codebook applied at scale using an LLM to tag qualitative social media posts. In the current wave of LLM-related research, a few studies have focused on methodologically validating how LLMs can be leveraged for qualitative research [17, 30, 80]. However, to our knowledge, few studies to date demonstrate an application of these tools to deepen understanding of a particular subject area. We believe that validating our codebook and making it available for use shows that in certain contexts, LLMs can be leveraged to deepen scientific research.

7 LIMITATIONS

In our methodology, we make certain choices which may impact the generalizability, representativeness, and nuance of our analysis. We chose to map Nextdoor neighborhoods onto a numerical gentrification index. While we lean on data provided by the UDP due to its contextual relevance and sensitivity to Atlanta, this choice can be reductive. For example, a few neighborhoods were not coded by UDP and are therefore left out. Additionally, we are likely missing further nuance by reducing the index to three types of gentrification. We make trade-offs for the sake of interpretibility, but acknowledge that these are not a substitute for more contextual, community-engaged action and knowledge. We chose to omit posts with fewer than five reactions, which may have biased the types of posts included in the analysis. While we believe this analysis thus prioritized the highest-impact community surveillance posts on Nextdoor, it may introduce bias; for example by favoring more incendiary posts that are likely to receive engagement. The MCA analysis is a tool which allows us to visualize and make sense of complex data in a specific way. However, the method does not allow us to include control variables such as population against income levels. Therefore, our quantitative findings rely on correlations and should be taken as contextually informative but not definitive. Nonetheless, we make every effort to show that the posts we label represent a reasonable sample of the high-impact posts made across Nextdoor neighborhoods in Atlanta.

8 CONCLUSION

While some have upheld local online spaces as spaces for civic local discussions, critics of Nextdoor highlight worrying trends of digital exclusion. Our study is one of the first empirical works to leverage data directly from Nextdoor to understand how the content of individual posts reflect larger patterns of community surveillance. Our findings suggest there are pockets of healthy discussion on Nextdoor – instances where neighbors warn each other of road hazards, or encourage each other to take action for the good of the community. However, we find evidence that these types of civil, “neighborly” conversations blossom most in already privileged neighborhoods, and thus are mostly beneficial to those who already have resources. At the same time, direct incident reports about crime are more prevalent, and often compromise the privacy of those who are seen as transgressing in a neighborhood, calling simultaneously for general vigilance from the community. We observe that Nextdoor can become a tool for further othering and gentrification. Through user-generated posts, Nextdoor can facilitate the development of a digital neighborhood constitution that cements who is included and who is not.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. (SaTC:2120651). We are grateful to the Harvard Cyber Law Clinic for helping us gauge how to conduct ethical data collection, specifically Christopher Bavitz and his students. A big thanks to Megan Brown for sharing code without which this work would not have been possible. We also thank the Citizens and Technology Lab, the Social Technologies Lab, and the Digital Life Initiative at Cornell University for constructive feedback throughout the study design and analysis. Finally, we thank our colleagues Eugene Bagdasaryan, Ero Balsa, A.F. Cooper, Ilan Mandel, Emily Tseng, Helen Nissenbaum and Pheobe Sengers.

REFERENCES

[1] Hervé Abdi and Dominique Valentin. 2007. Multiple correspondence analysis. *Encyclopedia of measurement and statistics* 2, 4 (2007), 651–657.

[2] Tyler Adkisson. 2022. How Did Nextdoor Get So Popular? *Scripps News* (2022). https://scrippsnews.com/stories/how-did-nextdoor-get-so-popular/

[3] Anmol Agarwal, Shrey Gupta, Vamshi Bonagiri, Manas Gaur, Joseph Reagle, and Ponnurangam Kumaraguru. 2023. Towards Effective Paraphrasing for Information Disguise. In *European Conference on Information Retrieval*. Springer, 331–340. https://doi.org/10.1007/978-3-031-28238-6_22

[4] Marianne Aubin Le Quéré, Ting-Wei Chiang, Karen Levy, and Mor Naaman. 2022. Information Needs of Essential Workers During the COVID-19 Pandemic. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–19. https://doi.org/10.1145/3555197

[5] Marianne Aubin Le Quéré, Mor Naaman, and Jenna Fields. 2024. Local, Social, and Online: Comparing the Perceptions and Impact of Local Online Groups and Local Media Pages on Facebook. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024). https://doi.org/10.1145/3637377

[6] Marianne Aubin Le Quéré, Hope Schroeder, Casey Randazzo, Jie Gao, Ziv Epstein, Simon Perrault, David Mimno, Louise Barkhuus, and Hanlin Li. 2024. LLMs as Research Tools: Applications and Evaluations in HCI Data Work. In *Proceedings of the 2024 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. https://doi.org/10.1145/3613905.3636301

[7] Swati Bhatt. 2017. *The Private World of Sharing and Cooperation*. Springer International Publishing, Cham, 119–132. https://doi.org/10.1007/978-3-319-47250-8_7

[8] Stefano Bloch. 2022. Aversive racism and community-instigated policing: The spatial politics of Nextdoor. *Environment and Planning C: Politics and Space* 40, 1 (2022), 260–278. https://doi.org/10.1177/23996544211019754

[9] Ian Bogost. 2018. What Petty Nextdoor Posts Reveal About America. *The Atlantic* (Jun 2018). https://www.theatlantic.com/magazine/archive/2018/07/nextdoor-american-communities/561746/

[10] Simone Browne. 2015. *Dark matters: On the surveillance of blackness*. Duke University Press.

[11] Kelly Caine. 2016. Local Standards for Sample Size at CHI *(CHI ’16)*. Association for Computing Machinery, New York, NY, USA, 981–992. https://doi.org/10.1145/2858036.2858498

[12] Manuel Castells. 2002. *The Internet Galaxy: Reflections on the Internet, Business, and Society*. OUP Oxford.

[13] Vania Ceccato. 2019. Eyes and apps on the streets: From surveillance to sousveillance using smartphones. *Criminal Justice Review* 44, 1 (2019), 25–41. https://doi.org/10.1177/0734016818818696

[14] Alissa Centivany and Bobby Glushko. 2016. "Popcorn Tastes Good" Participatory Policymaking and Reddit’s. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. ACM, 1126–1137. https://doi.org/10.1145/2858036.2858516

[15] Janet Chan. 2008. The new lateral surveillance and a culture of suspicion. In *Surveillance and governance: Crime control and beyond*. Vol. 10. Emerald Group Publishing Limited, 223–239. https://doi.org/10.1016/S1521-6136(07)00210-2

[16] Janet X Chen, Allison McDonald, Yixin Zou, Emily Tseng, Kevin A Roundy, Acar Tamersoy, Florian Schaub, Thomas Ristenpart, and Nicola Dell. 2022. Trauma-informed computing: Towards safer technology experiences for all. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–20. https://doi.org/10.1145/3491102.3517475

[17] Robert Chew, John Bollenbacher, Michael Wenger, Jessica Speer, and Annice Kim. 2023. LLM-Assisted Content Analysis: Using Large Language Models to Support Deductive Coding. *arXiv preprint arXiv:2306.14924* (2023). https://doi.org/10.48550/arXiv.2306.14924

[18] Ishita Chordia, Lena-Phuong Tran, Tala June Tayebi, Emily Parrish, Sheena Erete, Jason Yip, and Alexis Hiniker. 2023. Deceptive Design Patterns in Safety Technologies: A Case Study of the Citizen App. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI ’23)*. Association for Computing Machinery, New York, NY, USA, Article 193, 18 pages. https://doi.org/10.1145/3544548.3581258

[19] Eric Corbett and Yanni Loukissas. 2019. Engaging Gentrification as a Social Justice Issue in HCI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI ’19)*. Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3290605.3300510

[20] Patrício Soares Costa, Nadine Correia Santos, Pedro Cunha, Jorge Cotter, Nuno Sousa, et al. 2013. The use of multiple correspondence analysis to explore associations between categories of qualitative variables in healthy ageing. *Journal of aging research* 2013 (2013). https://doi.org/10.1155/2013/302163

[21] Anubrata Das, Brandon Dang, and Matthew Lease. 2020. Fast, accurate, and healthier: Interactive blurring helps moderators reduce exposure to harmful content. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, Vol. 8. 33–42. https://doi.org/10.1609/hcomp.v8i1.7461

[22] Í de Macedo Bernardino, LM Santos, AVP Ferreira, TLM de Almeida Lima, LM da Nóbrega, and S d’Avila. 2018. Multiple correspondence analysis as a strategy to explore the association between categories of qualitative variables related to oral–maxillofacial trauma and violent crimes at the community level. *International journal of oral and maxillofacial surgery* 47, 3 (2018), 339–344. https://doi.org/10.1016/j.ijom.2017.08.001

[23] Jonas De Meulenaere, Cédric Courtois, and Koen Ponnet. 2020. Bottom-up hyperlocal media in Belgium: Facebook-groups as collaborative neighborhood awareness systems. In *The Routledge Companion to Local Media and Journalism*. 398–407. https://doi.org/10.4324/9781351239943

[24] Catherine D'Ignazio and Lauren Klein. 2020. 3. on rational, scientific, objective viewpoints from mythical, imaginary, impossible standpoints. *Data Feminism* (2020).

[25] Sheena Erete, Love Nicole, Jesse Mumm, Anfal Boussayoud, and Ihudiya Finda Ogbonnaya-Ogburu. 2016. "That Neighborhood is Sketchy!": Examining Online Conversations about Social Disorder in Transitioning Neighborhoods. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 1180–1186. https://doi.org/10.1145/2851581.2892550

[26] Sheena Erete, Love Nicole, Jesse Mumm, Anfal Boussayoud, and Ihudiya Finda Ogbonnaya-Ogburu. 2016. "That Neighborhood is Sketchy!": Examining Online Conversations about Social Disorder in Transitioning Neighborhoods. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 1180–1186. https://doi.org/10.1145/2851581.2892550

[27] Casey Fiesler, Jialun Jiang, Joshua McCann, Kyle Frye, and Jed Brubaker. 2018. Reddit rules! characterizing an ecosystem of governance. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 12. https://doi.org/10.1609/icwsm.v12i1.15033

[28] Dídac Florensa, Pere Godoy, Jordi Mateo, Francesc Solsona, Tere Pedrol, Miquel Mesas, and Ramon Pinol. 2021. The use of multiple correspondence analysis to explore associations between categories of qualitative variables and cancer incidence. *IEEE Journal of Biomedical and Health Informatics* 25, 9 (2021), 3659–3667. https://doi.org/10.1109/JBHI.2021.3073605

[29] Michel Foucault. 1975. *Discipline and Punish*. Vintage.

[30] Jie Gao, Yuchen Guo, Gionnieve Lim, Tianqin Zhan, Zheng Zhang, Toby Jia-Jun Li, and Simon Tangi Perrault. 2023. CollabCoder: A GPT-Powered Workflow for Collaborative Qualitative Analysis. *arXiv preprint arXiv:2304.07366* (2023). https://doi.org/10.1145/3584931.3607500

[31] Fabrizio Gilardi, Meysam Alizadeh, and Maël Kubli. 2023. Chatgpt outperforms crowd-workers for text-annotation tasks. *arXiv preprint arXiv:2303.15056* (2023). https://doi.org/10.1073/pnas.2305016120

[32] Oliver L Haimson and Anna Lauren Hoffmann. 2016. Constructing and enforcing" authentic" identity online: Facebook, real names, and non-normative identities. *First Monday* (2016). https://doi.org/10.5210/fm.v21i6.6791

[33] Katherine B Hankins, Robert Cochran, and Kate Driscoll Derickson. 2012. Making space, making race: Reconstituting white privilege in Buckhead, Atlanta. *Social & Cultural Geography* 13, 4 (2012), 379–397. https://doi.org/10.1080/14649365.2012.688851

[34] Wanrong He, Mitchell L Gordon, Lindsay Popowski, and Michael S Bernstein. 2023. Cura: Curation at Social Media Scale. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–33. https://doi.org/10.1145/3610186

[35] Jessica M Hill, Ruth Jobling, Thomas V Pollet, and Daniel Nettle. 2014. Social capital across urban neighborhoods: A comparison of self-report and observational data. *Evolutionary Behavioral Sciences* 8, 2 (2014), 59. https://doi.org/10.1037/h0099131

[36] Sarah Holder. 2020. *Are you a robot?* https://www.bloomberg.com/news/articles/2020-02-12/what-police-departments-get-from-using-nextdoor

[37] Sarah Holder and Fola Akinnibi. 2022. *Nextdoor's Quest to Beat Toxic Content and Make Money*. Retrieved September 10, 2023 from https://www.bloomberg.com/news/features/2022-06-09/nextdoor-s-quest-to-beat-toxic-content-and-make-money

[38] Waleed Iqbal, Vahid Ghafouri, Gareth Tyson, Guillermo Suarez-Tangil, and Ignacio Castro. 2023. Lady and the Tramp Nextdoor: Online Manifestations of Economic Inequalities in the Nextdoor Social Network. https://doi.org/10.48550/arXiv.2304.05232 arXiv:2304.05232 [cs.SI]

[39] Waleed Iqbal, Gareth Tyson, Vahid Ghafouri, Guillermo Suarez-Tangil, and Ignacio Castro. 2022. Exploring online manifestations of real-world inequalities. In *Proceedings of the 22nd ACM Internet Measurement Conference*. 762–763. https://doi.org/10.1145/3517745.3563027

[40] Nicholas W Jankowski, Martine Van Selm, and Ed Hollander. 2005. On crafting a study of digital community networks. *Community Informatics: Shaping Computer-Mediated Social Networks* 101 (2005).

[41] Chapple K., Thomas T., and Zuk M. 2021. *Urban Displacement Project Website*. Retrieved September 10, 2023 from https://www.urbandisplacement.org/about/what-we-do/

[42] Cristina Kadar, Yiea-Funk Te, Raquel Rosés Brüngger, and Irena Pletikosa Cvijikj. 2016. Digital Neighborhood Watch: To share or not to share?. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 2148–2155. https://doi.org/10.1145/2441776.2441853

[43] Larry Keating. 2001. *Atlanta: Race, class and urban expansion*. Vol. 8. Temple University Press.

[44] Liam Kennedy and Madelaine Coelho. 2022. Security, Suspicion, and Surveillance? There's an App for That. *Surveillance & Society* 20, 2 (2022), 127–141. https://doi.org/10.24908/ss.v20i2.14536

[45] Julilly Kohler-Hausmann. 2007. The crime of survival": Fraud prosecutions, community surveillance, and the original" welfare queen. *Journal of Social History* (2007), 329–354. https://doi.org/10.1353/jsh.2008.0002

[46] Jerry Kudlats, Arthur Money, and Joseph F Hair Jr. 2014. Correspondence analysis: A promising technique to interpret qualitative data in family business research. *Journal of Family Business Strategy* 5, 1 (2014), 30–40. https://doi.org/10.1016/j.jfbs.2014.01.005

[47] Rahim Kurwa. 2019. Building the Digitally Gated Community: The Case of Nextdoor. *Surveillance & Society* 17, 1/2 (2019), 111–117. https://doi.org/10.24908/ss.v17i1/2.12927

[48] Sarah Lacy. 2013. *Racial Profiling Via Nextdoor.com*. Retrieved September 10, 2023 from https://web.archive.org/web/20150312061336/https://pando.com/2013/01/22/nextdoors-unexpected-killer-use-case-crime-and-safety

[49] Katie Lambright. 2019. Digital redlining: The Nextdoor app and the neighborhood of make-believe. *Cultural Critique* 103 (2019), 84–90. https://doi.org/10.1353/cul.2019.0023

[50] James Ledvinka. 1972. The intrusion of race: Black responses to the white observer. *Social Science Quarterly* (1972), 907–920.

[51] Jenny Lee and Chloe Ahn. 2023. Colonizers in the neighborhood: a critical discourse analysis of Nextdoor users' postracial strategies. *Information, Communication & Society* (2023), 1–16. https://doi.org/10.1080/1369118X.2023.2252484

[52] Sam Levin. 2015. *Racial Profiling Via Nextdoor.com*. Retrieved September 10, 2023 from https://eastbayexpress.com/racial-profiling-via-nextdoorcom-2-1/

[53] Sheena Lewis and Dan A. Lewis. 2012. Examining Technology That Supports Community Policing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 1371–1380. https://doi.org/10.1145/2207676.2208595

[54] Claudia López and Rosta Farzan. 2015. Lend me sugar, I am your neighbor! A content analysis of online forums for local communities. In *Proceedings of the 7th International Conference on Communities and Technologies*. 59–67.

[55] Claudia López, Rosta Farzan, and Yu-Ru Lin. 2017. Behind the myths of citizen participation: Identifying sustainability factors of hyper-local information systems. *ACM Transactions on Internet Technology (TOIT)* 18, 1 (2017), 1–28. https://doi.org/10.1145/3093892

[56] Claudia A López and Brian S Butler. 2013. Consequences of content diversity for online public spaces for local communities. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 673–682. https://doi.org/10.1145/2441776.2441851

[57] David Lyon. 1994. *Electronic Eye: The Rise of Surveillance Society*. University of Minnesota Press. http://ebookcentral.proquest.com/lib/columbia/detail.action?docID=310274

[58] David Lyon. 2003. *Surveillance as social sorting: privacy, risk, and digital discrimination*. Routledge.

[59] Christina A. Masden, Catherine Grevet, Rebecca E. Grinter, Eric Gilbert, and W. Keith Edwards. 2014. Tensions in Scaling-up Community Social Media: A Multi-Neighborhood Study of Nextdoor. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 3239–3248. https://doi.org/10.1145/2556288.2557319

[60] Richard M McGahey. 1986. Economic conditions, neighborhood organization, and urban crime. *Crime and justice* 8 (1986), 231–270. https://doi.org/10.1007/BF02900932

[61] Gaia Mosconi, Matthias Korn, Christian Reuter, Peter Tolmie, Maurizio Teli, and Volkmar Pipek. 2017. From Facebook to the Neighbourhood: Infrastructuring of Hybrid Community Engagement. *Computer Supported Cooperative Work (CSCW)* 26, 4 (2017), 959–1003. https://doi.org/10.1007/s10606-017-9291-z

[62] NextDoor. 2023. *Nextdoor, an app for neighborhoods where you can get local tips, buy and sell items, and more.* Retrieved June 5, 2023 from https://nextdoor.com/

[63] Étienne Ollion, Rubing Shen, Ana Macanovic, and Arnault Chatelain. 2024. The dangers of using proprietary LLMs for research. *Nature Machine Intelligence* (2024), 1–2. https://doi.org/10.1038/s42256-023-00783-6

[64] OpenAI. 2023. GPT-4 Technical Report. arXiv:2303.08774 [cs.CL]

[65] Will Payne. 2017. Welcome to the Polygon: Contested Digital Neighborhoods and Spatialized Segregation on Nextdoor. *Computational Culture* (2017). http://computationalculture.net/welcome-to-the-polygon-contested-digital-neighborhoods-and-spatialized-segregation-on-nextdoor/

[66] Pew Research Centre. 2019. For Local News, Americans Embrace Digital but Still Want Strong Community Connection. https://www.pewresearch.org/journalism/2019/03/26/for-local-news-americans-embrace-digital-but-still-want-strong-community-connection/

[67] Rafael Prieto Curiel, Stefano Cresci, Cristina Ioana Muntean, and Steven Richard Bishop. 2020. Crime and its fear in social media. *Palgrave Communications* 6, 1 (2020). https://doi.org/10.1057/s41599-020-0430-7

[68] Anaïk Purenne and Grégoire Palierse. 2017. Towards cities of informers? Community-based surveillance in France and Canada. *Surveillance & Society* 15, 1 (2017), 79–93. https://doi.org/10.24908/ss.v15i1.5619
[69] Robert D Putnam. 2000. Bowling alone: America's declining social capital. In *Culture and politics: A reader*. Springer, 223–234.
[70] Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084* (2019). https://doi.org/10.48550/arXiv.1908.10084
[71] Dennis P Rosenbaum. 1988. Community crime prevention: A review and synthesis of the literature. *Justice Quarterly* 5, 3 (1988), 323–395. https://doi.org/10.1080/07418828800089781
[72] Sarah Fioroni. 2022. Americans' Social Networks and Social Media, Not Traditional News Sources, Drive Community Engagement. https://knightfoundation.org/articles/americans-social-networks-and-social-media-not-traditional-news-sources-drive-community-engagement/
[73] The Markup Staff. 2021. *Why Web Scraping Is Vital to Democracy*. Retrieved December 10, 2023 from https://themarkup.org/news/2020/12/03/why-web-scraping-is-vital-to-democracy
[74] Henri Tajfel, Michael G Billig, Robert P Bundy, and Claude Flament. 1971. Social categorization and intergroup behaviour. *European journal of social psychology* 1, 2 (1971), 149–178. https://doi.org/10.1002/ejsp.2420010202
[75] Ori Tenenboim and Akiba A Cohen. 2015. What prompts users to click and comment: A longitudinal study of online news. *Journalism* 16, 2 (2015), 198–217. https://doi.org/10.1177/1464884913513996
[76] Petter Törnberg. 2023. How to use LLMs for Text Analysis. *arXiv preprint arXiv:2307.13106* (2023). https://doi.org/10.48550/arXiv.2307.13106
[77] Katherine Van Koevering, Yiquan Hong, and Jon Kleinberg. 2022. Exporting Geography Into A Virtual Landscape: A Global Pandemic Locally Discussed. *arXiv preprint arXiv:2210.07187* (2022). https://doi.org/doi.org/10.48550/arXiv.2210.07187
[78] Kaveh Waddell. 2020. *The Police Officer 'Nextdoor'*. https://www.theatlantic.com/technology/archive/2016/05/nextdoor-social-network-police-seattle/481164/
[79] Qunfang Wu, Louisa Kayah Williams, Ellen Simpson, and Bryan Semaan. 2022. Conversations About Crime: Re-Enforcing and Fighting Against Platformed Racism on Reddit. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–38.
[80] Ziang Xiao, Xingdi Yuan, Q Vera Liao, Rania Abdelghani, and Pierre-Yves Oudeyer. 2023. Supporting Qualitative Analysis with Large Language Models: Combining Codebook with GPT-3 for Deductive Coding. In *Companion Proceedings of the 28th International Conference on Intelligent User Interfaces*. 75–78. https://doi.org/10.1145/3581754.3584136
[81] Shoshana Zuboff. 2019. *The Age of Surveillance Capitalism*. PublicAffairs.

A CRIME-RELATED KEYWORDS

To generate the keywords, three researchers on this project independently coded 300 randomly selected Nextdoor posts from the larger dataset. The researchers labeled each post for whether it was relevant to community surveillance, and listed three keywords for each relevant post (e.g. "cars," "speeding," "cops"). To limit bias stemming from researcher interpretation at this stage, keywords noted by the researchers were drawn verbatim from the post being coded. To ensure comprehensiveness, we additionally include a pre-validated list of crime-related keywords on social media [67]. We note that, while the original list relied on verbatim matching and thus separated between root and variations of the same word, S-BERT is less likely to be affected by such differences and thus only the root forms of words were included in our list; e.g. "steal" and "stealing" were not both added as keywords. The research team discussed and aggregated initial keywords into a set of 47 community surveillance keywords. The keywords were subsequently narrowed by eliminating those that were less relevant to the corpus (few matching posts), or where post-keyword occurrences had a pearson coefficient that was 0.5 or greater with another keyword (e.g. "cops" and "police"). When two keywords were found to be highly correlated, we retained the keyword that matched the most posts in our corpus. This process yielded 17 community surveillance-related keywords.

We include here the list of 17 community surveillance keywords:
`'property damage', 'police', 'safety', 'theft', 'noise', 'incident', 'vigilance', 'creep', 'violent crime', 'gun', 'victim', 'suspect', 'trespassing', 'enforcement', 'stalking', 'speeding', 'hitting', 'scary'`

We narrowed the keywords down to avoid those which were too highly correlated with each other. We include a correlation matrix of the keywords in Figure 6. We also show how many posts in our dataset matched each keyword in Figure 7.

B DATA VALIDATION

To examine how effective each of the post filtering decisions was for identifying community surveillance posts, we validate each step. This process was done through randomly sampling 100 posts using the different criteria, and then annotating each post with if it should be considered a community surveillance post. The community surveillance post occurrence rate for 100 randomly sampled posts using only the 10 words or more criteria (9% post inclusion rate), only the five-or-more reactions criteria (26% post inclusion rate), and only the BERT-based filtering criteria (50% post inclusion rate). We thus demonstrate that combining these strategies together (76% inclusion rate) was effective for isolating a large dataset of community surveillance posts. Though we do not claim that every community surveillance post is included in our final dataset, we collected a sufficient number of instances to be able to identify patterns and construct a typology, which is the main goal of this research [11].

C CODEBOOK

We include our final codebook and all interrater correlations in Table 5.

D CORRELATIONS

We map the Pearson correlations between the binary codebook variables in Figure 8. There is a cluster of positive correlations between posts that explicitly promote vigilance, provide a physical description of someone, and share a personal experience. Additionally, soliciting information or action is negatively correlated with expressing a personal opinion or explicitly calling for vigilance. Finally, posts that share a physical description negatively correlate with those that include a personal opinion.

We also show the correlation between different community concerns and community tactics in Figure 9. Physical description and vigilance are both quite strongly correlated with the criminal or suspicious person concern. Describing a personal experience and vigilance are both negatively correlated with posts tagged "not applicable."

E MCA

We include the Scree plot for the MCA in Figure 10. Figure 11 visually displays how the gentrification indices map onto the MCA dimensions. Figure 12a shows how census variables map onto the MCA dimensions. Figure 12b shows how supplementary platform variables map onto the MCA dimensions.

Table 5: Final codebook, code descriptions, and interrater agreement measured using unweighted Cohen's Kappa for all categories across 100 community surveillance posts. We accept the ratings provided by GPT-4 when they have an average reliability of 0.6 or higher, indicating at least moderate agreement. Only the codes *providing information* and *object* do not meet our required threshold.

Category	Code	Human 1 x Human 2 κ	Human 1 x AI κ	Human 2 x AI κ	Avg AI κ
Main post topic The central matter of interest in a post.	police activity, guns or gunshots, property damage, noise, criminal or suspicious person, theft, unsafe driving, sexual violence or harassment, inanimate object or animal, not applicable	0.95	0.79	0.82	0.81
Roles A function assumed by a person posting on the Nextdoor platform.	community member, administrator, organizer, or moderator, not applicable	1	0.65	0.65	0.65
Providing Information Any post that gives new, potentially beneficial, information to the reader.	yes, no	0.93	0.49	0.57	0.53
Explicitly calling for vigilance A post that explicitly asks people to be cautious, watchful, or be on the lookout for someone	yes, no	0.96	0.66	0.70	0.68
Describing a personal experience Anyone who is sharing an event they witnessed first-hand or that happened to them.	yes, no	0.91	0.81	0.77	0.79
Expressing a personal opinion Someone sharing their viewpoint on something without being prompted or with the intent to convince others.	yes, no	0.85	0.71	0.68	0.70
Soliciting information or action Explicit requests for something, e.g. pictures of an event.	yes, no	0.94	0.73	0.73	0.73
Object The main thing or person that is discussed in a post.	person, group of people, not applicable	0.97	0.26	0.25	0.26
Physical Description The post describes how a person looks e.g. their race, age, or gender.	yes, no	0.80	0.80	0.68	0.72
Primary Sentiment The main emotion a post is likely to evoke in the reader.	positive, neutral, negative	0.88	0.74	0.67	0.71

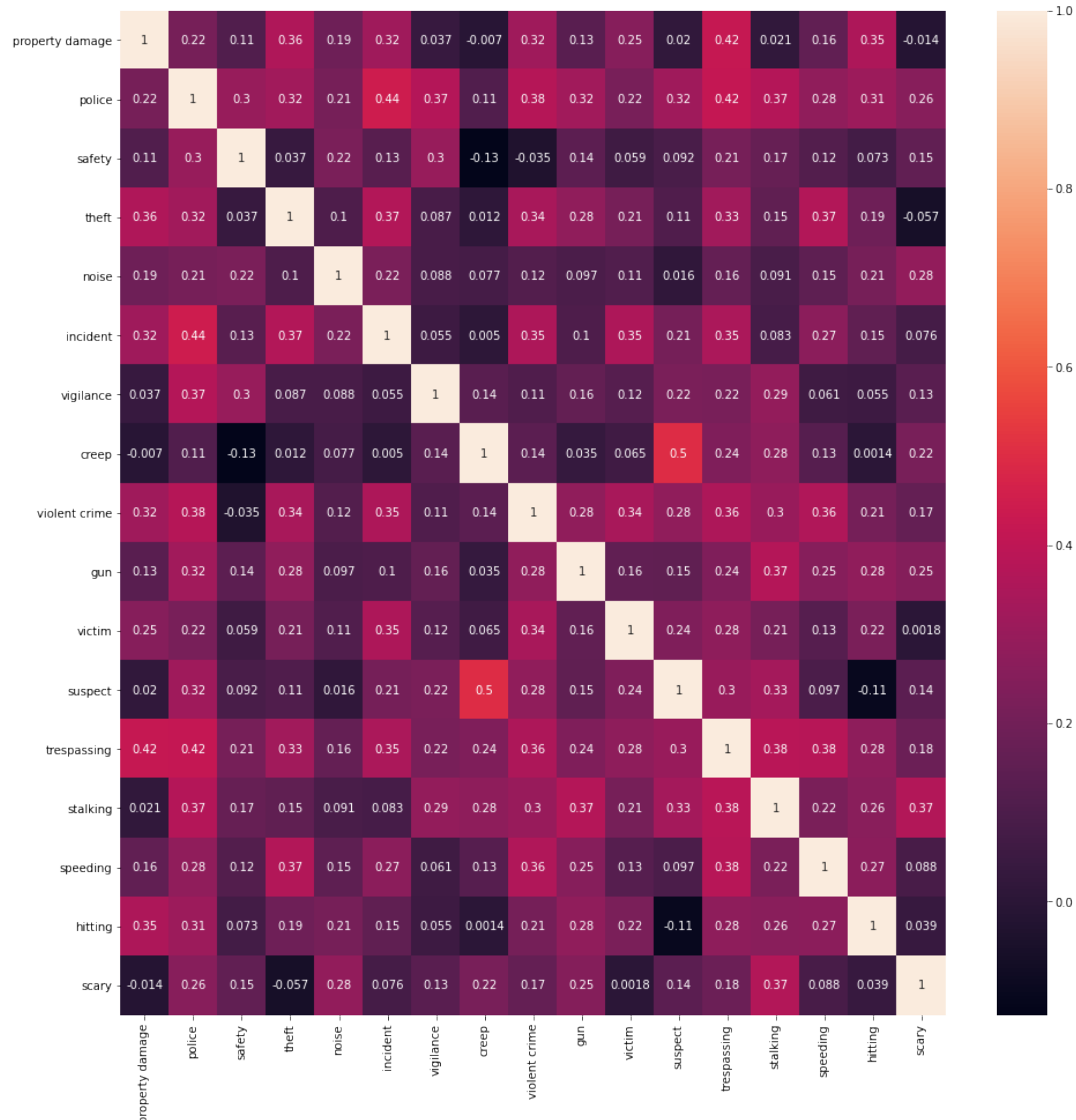


Figure 6: Correlation matrix of crime, suspicion, and surveillance keywords cosine similarities across the corpus. Keywords were removed if they had a higher than 0.5 correlation with another keyword or if they did not match many of the posts in the corpus.

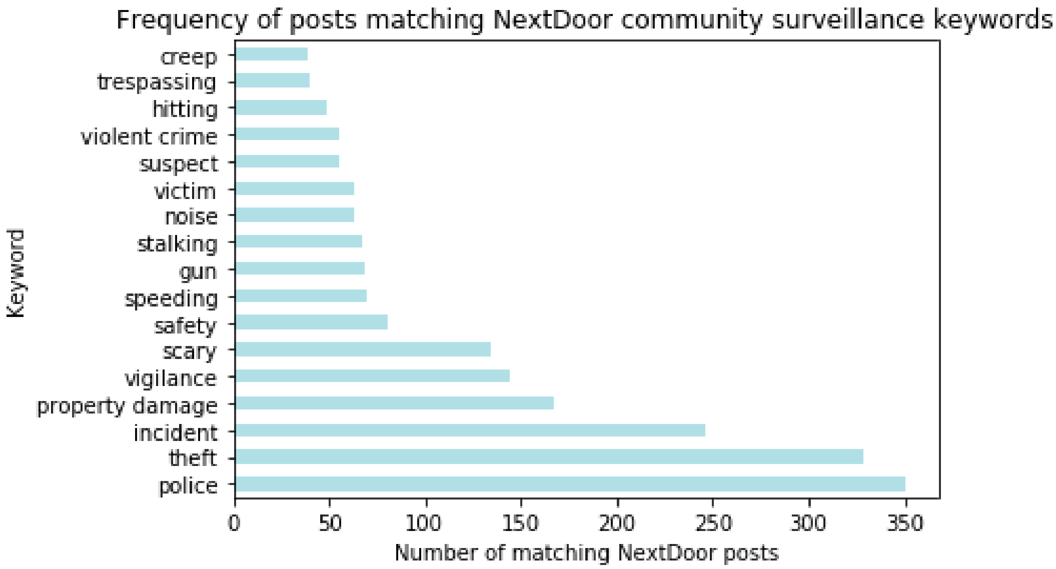


Figure 7: Figure depicts the distribution of Nextdoor posts that match one of the 17 keywords we have identified as relating to community surveillance. Posts matching the keywords *police, theft,* and *incident* are the most prevalent in our corpus.

Figure 8: Correlation plot depicts how the binary codes relate to each other. The strongest correlation is between between a poster publicly disclosing a physical description and explicitly calling for vigilance.

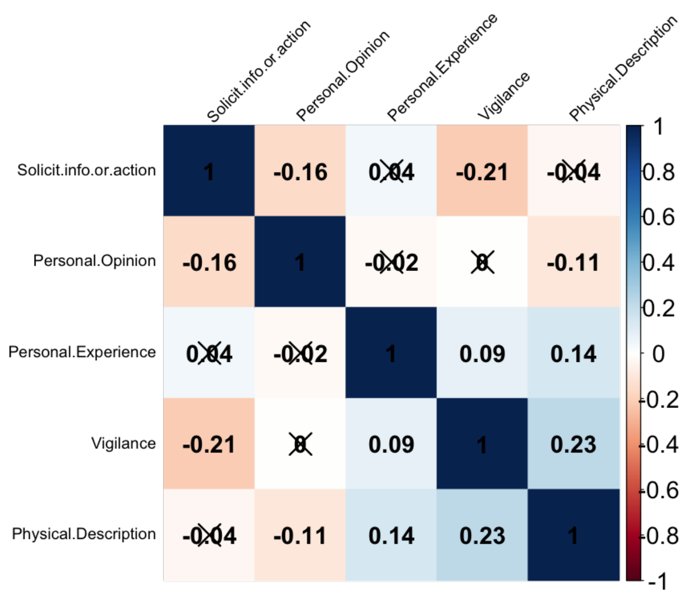
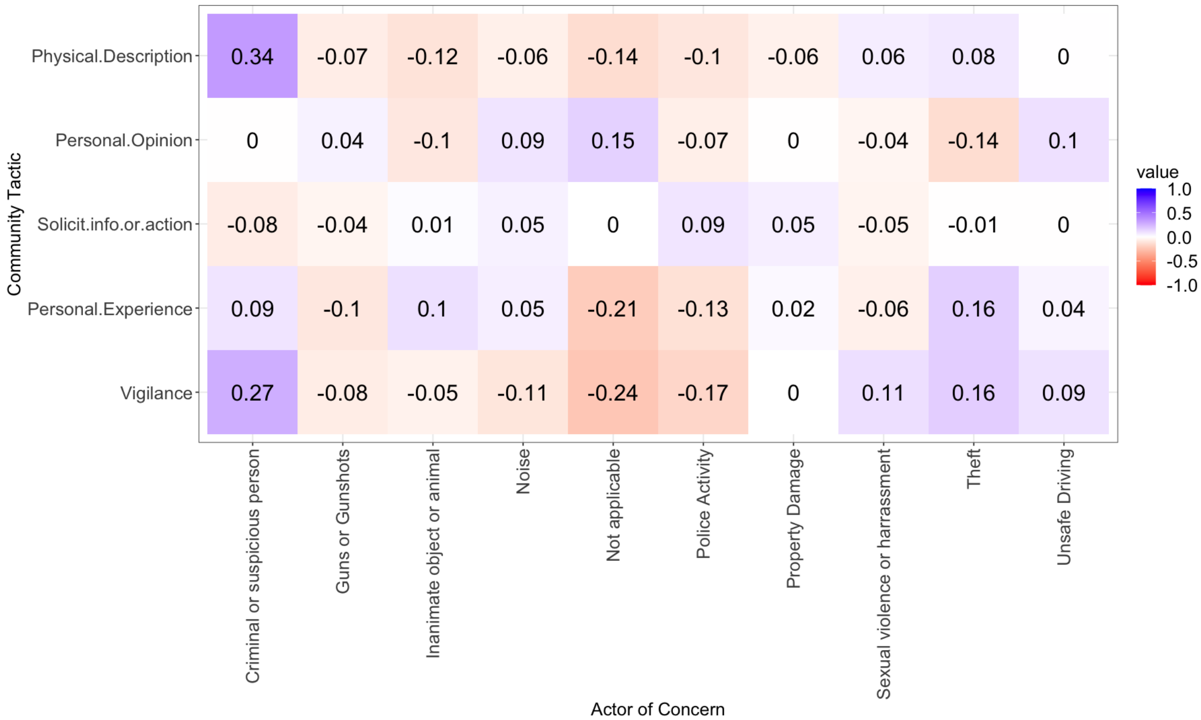


Figure 9: Heatmap depicts the community tactics (y-axis) most correlated with each of the actors of concern (x-axis).



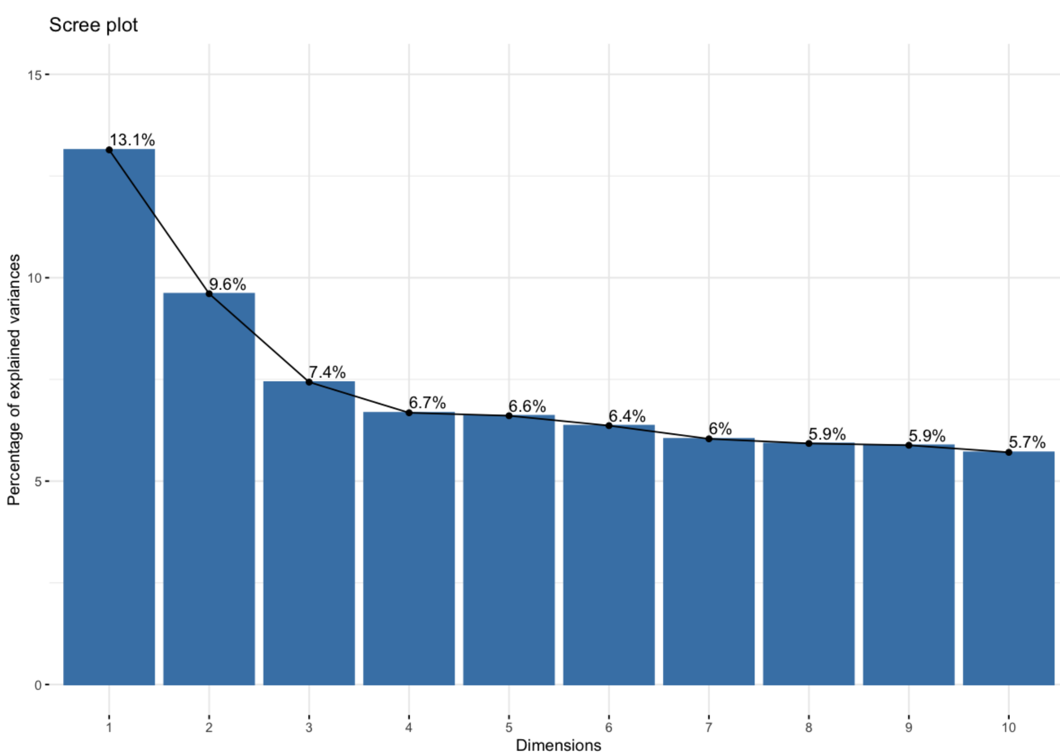


Figure 10: The scree plot describes how much inertia (variance) is described by each dimension. The two dimensions we focus on encapsulated about 23% of the total variance.

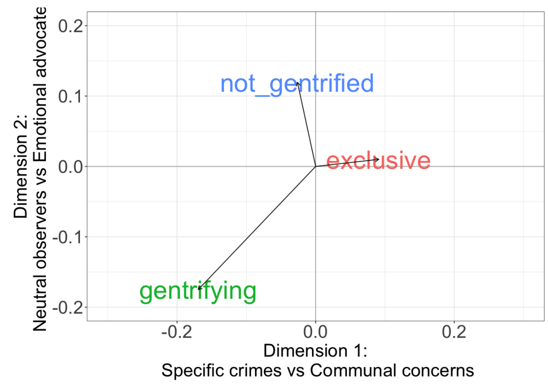
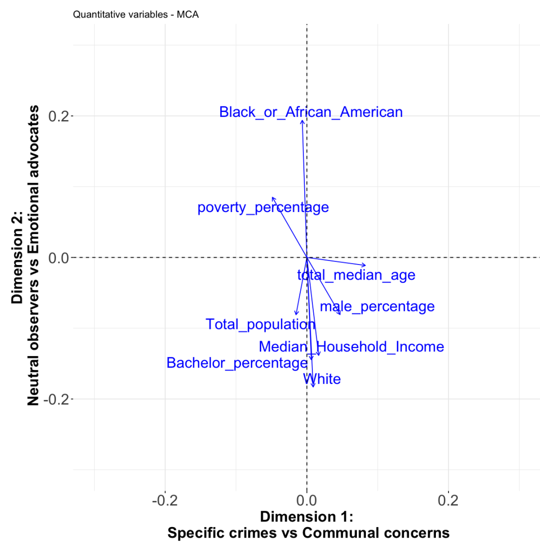
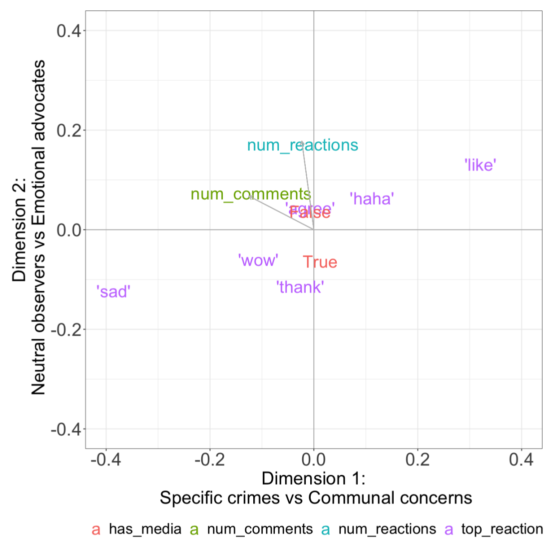


Figure 11: Visual representation of how the vectors for the three indices of gentrification map onto the two dimensions of the MCA.



(a) Census variables



(b) Platform variables

Figure 12: This figure shows various supplemental variables mapped to dimensions 1 and 2. On the left, we show how the census variables map onto the dimensions. On the right, various metadata from the platform, for example, the number of comments, are mapped on to the two dimensions.