Understanding People's Concerns and Attitudes Toward Smart Cities

Pardis Emami-Naeini
pardis@cs.duke.edu
Duke University
USA

Joseph Breda
joebreda@cs.washington.edu
University of Washington
USA

Wei Dai
weidai.d@tiktok.com
TikTok
USA

Tadayoshi Kohno
yoshi@cs.washington.edu
University of Washington
USA

Kim Laine
kim.laine@microsoft.com
Microsoft Research
USA

Shwetak Patel
shwetak@cs.washington.edu
University of Washington
USA

Franziska Roesner
franzi@cs.washington.edu
University of Washington
USA

ABSTRACT

Designing privacy-respecting and human-centric smart cities requires a careful investigation of people's attitudes and concerns toward city-wide data collection scenarios. To capture a holistic view, we carried out this investigation in two phases. We first surfaced people's understanding, concerns, and expectations toward smart city scenarios by conducting 21 semi-structured interviews with people in underserved communities. We complemented this in-depth qualitative study with a 348-participant online survey of the general population to quantify the significance of smart city factors (e.g., type of collected data) on attitudes and concerns. Depending on demographics, privacy and ethics were the two most common types of concerns among participants. We found the type of collected data to have the most and the retention time to have the least impact on participants' perceptions and concerns about smart cities. We highlight key takeaways and recommendations for city stakeholders to consider when designing inclusive and protective smart cities.

CCS CONCEPTS

• **Security and privacy** → **Social aspects of security and privacy**; **Usability in security and privacy**; • **Human-centered computing** → *Empirical studies in ubiquitous and mobile computing*.

KEYWORDS

Privacy, Ethics, Smart City, Internet of Things, IoT, Human Centered

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04…$15.00
https://doi.org/10.1145/3544548.3581558

ACM Reference Format:
Pardis Emami-Naeini, Joseph Breda, Wei Dai, Tadayoshi Kohno, Kim Laine, Shwetak Patel, and Franziska Roesner. 2023. Understanding People's Concerns and Attitudes Toward Smart Cities. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 24 pages. https://doi.org/10.1145/3544548.3581558

1 INTRODUCTION

Urbanization is rapidly increasing in countries around the world. For the first time in history [115], more than half of the world's population (55%) is living in the cities, and by 2050, this number is projected to increase to 68% [122]. This sharp growth in urbanization results in the rise of demands in areas including public safety, the environment, transportation, and energy. The concept of *smart city* has been introduced to satisfy these inevitable and increasing needs through the extensive use of information communication technologies (ICTs) and the Internet of Things (IoT). Although there is no universal consensus over what is or is not a smart city [24, 80], the use of ubiquitous sensor data collections to enhance the efficiency of urban operations has been frequently mentioned in the provided definitions [4]. For the purpose of this study, we use the Wikipedia definition of smart city [133] as "an urban area that uses different types of electronic methods and sensors to collect data. Insights gained from that data are used to manage assets, resources, and services efficiently."

Enabling the functionality of smart city technologies requires the deployed sensors to collect, store, and process large amounts of data from residents [103, 119]. City stakeholders and decision-makers deploy pervasive and costly data collections in neighborhoods with the hope and, in several cases, unjustified promise to improve the quality of life of their occupants. Hyper-local air quality sensors [120] to enhance health conditions, security cameras to enhance public safety [13], gunshot detectors to reduce gun violence [93], and people-counting sensors for efficient allocation of city resources [125] are only a few examples of common smart city projects in cities across the US [2].

Despite their advertised benefits, people are increasingly concerned about smart city IoT technologies in their neighborhoods, which in some cases have led to such projects being shut down or paused. The Sidewalk Toronto project [26, 63, 75] and smart street lights in San Diego [47, 118] made it clear that lack of timely attention to people's concerns toward the data collection and use of smart city sensors could be detrimental to smart city projects. To design equitable and privacy-respecting smart city technologies, we need to obtain an in-depth understanding of people's concerns, attitudes, and expectations toward city-wide IoT data collection and use scenarios. The current body of research on smart cities has generally taken a narrow and potentially dangerous view of smart cities as technical systems that require technical solutions. Indeed, very limited empirical research has been conducted to capture people's knowledge and attitudes toward smart cities [124], and none has looked into and surfaced the specific factors that impact people's concerns and expectations in a smart city.

While we are still at the early trial and error stages of smart city deployments, we, as researchers, city stakeholders, and policymakers can take responsible and proactive actions to mitigate the harms of these projects. To this end, we conducted a mixed-methods exploration of people's attitudes and concerns toward smart city data collection and use scenarios. We started our study by conducting 21 semi-structured interviews with participants living in Seattle, a city on the West Coast of the United States. Even in the same city, smart sensing projects are at different design and deployment stages, with some neighborhoods and communities experiencing these technologies earlier than others. To control for the confounding factor of the smart city deployment timeline, we specifically recruited participants from underserved neighborhoods that were identified by the City Council to have the most need for city resources and the highest priority in receiving smart sensing technologies in 10–15 years. These interviews enabled us to capture in-depth qualitative data on participants' concerns and knowledge of smart cities.

To be able to quantify the impact of smart city factors (e.g., data type, data access, data retention) on a broader population's preferences and concerns, we complemented our qualitative study with a large-scale online survey on Prolific, where we statistically analyzed data from 348 participants in the US. We aimed to investigate how concerns, attitudes, and preferences, which surfaced among underserved interview neighborhoods, manifest in the general population. Therefore, we did not limit our large-scale survey to specific cities, neighborhoods, or underserved demographics.

We found similar types of concerns and preferences between interview and survey participants. We especially surfaced two overarching types of concerns, namely *privacy* concerns, and *ethical* concerns. Among the tested smart city data collection and use scenarios, our interview and survey participants found the smart city vignettes involving *gunshot detectors* to have the most concerning ethical implications and the deployment of *security cameras* to be the most privacy-invasive.

Nevertheless, we identified a few differences between our interview and survey participants' perceptions toward smart city scenarios. Specifically, all of our interviewees were able to provide examples of how each of the presented smart city vignettes could potentially benefit and harm them and their neighborhoods. On the contrary, several of our survey participants struggled to identify potential benefits or harms of scenarios. This suggests that participants' perceived benefits and harms of smart city scenarios in underserved and low-income neighborhoods could be more pronounced than, and not easily generalizable to, the general population.

Our statistical analysis confirmed the aforementioned discrepancy. It specifically revealed that survey participants' level of income was significantly correlated with their reported type of concern. While the level of income did not have a statistically significant impact on participants' privacy concerns, lower-income respondents were significantly more concerned about the ethical implications of smart city scenarios when being deployed in their neighborhoods or cities. Some survey participants, in particular those with higher income levels, could hardly relate to certain smart city technologies; those with higher income attributed their struggle in identifying potential benefits or harms to not seeing any challenge (e.g., gun violence) in their neighborhoods or not seeing themselves in situations that could be impacted by such data collection technologies (e.g., gunshot detectors).

The contributions of our paper are three-fold:

- Through our qualitative data analysis, we surface participants' nuanced understanding and concerns toward smart city data collections.
- Through our quantitative data analysis, we measure the impact of various aspects of smart city data collections along with demographic factors on participants' concerns and attitudes.
- Grounded in qualitative and quantitative findings, we highlight several key privacy and ethical issues that various stakeholders should focus on to design equitable and privacy-respecting smart city technologies.

2 BACKGROUND AND RELATED WORK

We start this section by briefly talking about the concept of smart cities. We then explain how smart cities could pose a wide range of privacy and ethical harms to the public. We conclude this section by discussing the prior work on understanding the factors that influence people's attitudes toward smart city technologies and explaining how our research would fill the current gap in the smart city literature.

2.1 Smart Cities

The technological improvements and increased challenges caused by rapid urbanization have given rise to the concept of smart cities. Some of the cities around the world that are actively pursuing the philosophy of smart cities are London, Singapore, Amsterdam, Toronto, New York, Seattle, and Tokyo [56, 70, 72]. Although the journey of smart cities started in the 1990s [53], there is still no consensus over the definition of smart cities [4, 46, 98, 100]. The concept of the smart city refers to using information communication technologies (ICTs) and other methods to improve the quality of life of city occupants through enhancing the efficiency of city-wide services and operations [29, 60, 79, 98, 135, 136]. Internet of Things (IoT) is the backbone and core element of smart cities [98]. The IoT framework in smart cities encompasses a wide variety of sensors to enable components including security and surveillance, smart transportation, smart home devices, telemedicine, and ubiquitous and smart personal devices [98].

To facilitate the functionality of technologies in smart cities, city-wide sensors collect and process massive amounts of data and share them with various stakeholders (e.g., device manufacturers, law enforcement) [74]. Such ubiquitous data collections bring numerous benefits to the city and its inhabitants, such as enhanced public safety [13], improved air quality monitoring [120], and reduced traffic congestion [98]. However, smart city pervasive data collections could introduce the city residents to consequential risks and make them concerned [104, 108], which have not been thoroughly explored.

2.2 Privacy and Ethical Risks in Smart Cities

Privacy can be defined as a right to be free from unwanted intrusions [97], and this right has been frequently violated by the vision for automation and control in smart city projects [12, 50, 64]. Highlighting the massive data collection and lack of informed consent [22], researchers have identified several privacy threats in smart cities, including user profiling and location tacking [138].

In addition to privacy concerns, prior work has specified surveillance as one of the primary ethical challenges of emerging smart city technologies [65, 77, 87]. An example of such violation is racial biases in face recognition systems [25]. Another problematic example of city-wide IoT technologies is the use of gunshot detectors placed on streetlights across different cities throughout the US, including Chicago, Sacramento, and Philadelphia. Despite being deployed to enhance the safety of neighborhoods, these technologies have been shown to have little impact on preventing crimes [40] and, in several cases, being harmful to marginalized communities [96].

Moreover, privacy advocates and policymakers have criticized the data management practices of smart city projects. A few of the high-profile cases of privacy violations of smart city projects were in New York, Los Angeles, San Diego, and Toronto. In New York City, privacy experts raised concerns about LinkNYC Kiosks [83], which were mainly purposed to provide free Wi-Fi for the public around the city. Experts were concerned as people's collected data were being sold to third parties, and there was no adequate transparency over the camera activation of the kiosks [105]. In another example, American Civil Liberties Union (ACLU) filed a lawsuit against Los Angeles over their demand for having real-time location data from scooter rental companies [62].

People's privacy and security concerns in smart cities could impede the adoption of smart city technologies [90, 101, 129]. Some smart city projects have been shut down, partly due to increased privacy concerns. For example, in San Diego, the smart streetlights, which were equipped with sensors such as air quality sensors, cameras, and microphones, faced massive push-back from privacy advocates. As it turned out, the video footage from the streetlights was used solely as a surveillance tool by the police department to solve crimes for months without being disclosed to the city council or the public [92, 118]. This project has been recently shut down [47], and the city government is working on implementing new privacy policies [48]. In another example, Sidewalk Labs stopped their smart city project in Toronto [63]. Since the launch of their project, they have faced public uproar toward their concerning privacy practices. Specifically, the project experienced continuous backlash from privacy advocates over its lack of transparency and accountability around the collection and use of people's data in public locations [26, 75].

2.3 People's Attitudes and Concerns toward Smart City Technologies

Due to their significant impact on the success of smart city projects, researchers have acknowledged the importance of understanding people's perceived concerns and attitudes toward technologies in smart cities [14, 27, 31, 32, 123, 124]. For example, Belanche-Gracia et al. found that people's privacy concerns toward smart city technologies in Spain significantly influence their use and adoption of such technologies [14]. In addition, researchers have investigated the impact of privacy and ethical concerns on the success of crowdsourcing smart city projects [31, 32].

Limited research has been conducted to understand the factors that influence such concerns toward smart city IoT technologies. Van Heek et al. surfaced the impact of the location of smart city technologies on participants' privacy and security perceptions toward such data collections [123]. They found that surveillance technologies are more welcomed in city locations with higher crime rates (e.g., train stations), where participants were concerned about their safety, compared to places that were perceived to be safer.

Various context-related factors could potentially impact people's concerns and perceptions toward IoT data collection scenarios. Among other factors, Lee and Kobsa found that participants' level of concern toward data collection of IoT technologies in buildings greatly depends on who has access to the collected data [81, 82]. In the context of consumer IoT devices, Emami-Naeini et al. found that participants are concerned about the security and privacy practices of smart home devices, including the type of data they collect and the retention time [43].

In the smart city context, Zoonen et al. proposed a privacy framework for researchers and others who are interested in studying people's privacy concerns in a smart city. The proposed framework considers three dimensions, namely, the type of collected data, the purpose of data collection, and access to the collected data [124]. As acknowledged by Zoonen et al. and to the best of our knowledge, no empirical research has been conducted to understand the impact of various dimensions of IoT data collection factors on people's concerns, attitudes, and expectations.

Our work complements this line of research by providing a deep understanding of people's concerns and preferences in a smart city and quantifying how different factors impact such attitudes and concerns. Through rich qualitative data and statistical analysis of large-scale survey responses, we study the impact of three data collection factors of `data_type`, `data_access`, and `data_retention`, alongside demographic factors (e.g., level of income), on participants' concerns and assessments of smart city data collection and use scenarios.

Factor	Description	Levels
`data_type`	The type of data that is being collected by the sensor	Real-time video footage from people Presence of gunshots Number of people Air quality
`data_access`	Who has access to the collected data	Everybody Law enforcement officers Mayor's office Insurance companies
`data_retention`	For how long the collected data will be retained	One day Forever

Table 1: The factors and their levels that we varied and presented in smart city vignettes.

3 METHODS

We conducted a 21-participant semi-structured interview study followed by a 348-participant survey study on the Prolific crowdsourcing platform. The goal of the interview study was to capture an in-depth understanding of participants' perceptions of a smart city. We complemented this rich qualitative knowledge with a large-scale online survey to numerically quantify the impact of various demographic and data collection factors on concerns and attitudes toward smart city data collection scenarios.

The complete lists of interview and survey questions are provided in Appendices A and B, respectively. The study protocol was approved by our Institutional Review Board (IRB) and we obtained informed consent from all interview and survey participants.

3.1 Semi-Structured Interview Study

Recruitment and compensation. In March 2021, we recruited 23 participants from Seattle, a large city on the West Coast of the United States. To have participants who are somewhat familiar with the concept of smart city, we selected a city in the US with advanced city-wide technologies, including a wide range of environmental sensors, surveillance cameras, smart traffic lights, and gunshot detectors. Due to health precautions during the COVID-19 Pandemic, all interviews were conducted remotely over the Zoom application. [1] Each interview took on average one hour to be completed. We compensated each participant with a $25 local grocery store gift card.

Based on several criteria, including access to city services and displacement risk, the city government identified the most underserved neighborhoods that were populated by marginalized communities. To improve the quality of life in those neighborhoods, the city plans to invest in and develop extensive sensing technologies over the next decade. We selected those neighborhoods and recruited a diverse sample of participants by posting flyers and using the Craigslist platform to advertise our study. All participants had to be at least 18 years old.

In the recruitment material, we advertised our study as an interview about the experience of living in the city, not to prompt participants about smart cities. Participants who were interested in our interview study had to answer a screening survey, where we asked about the neighborhoods they live in, as well as some demographic questions. We asked participants for their email addresses and used that to invite them to our interview study. We provide the complete list of screening questions in Appendix A.1.

[1]https://zoom.us/download

Interview questions. We structured the interview questions in two sections. We provide the list of interview questions in Appendix A.2.

(1) *Knowledge and awareness about a smart city:* We asked participants to explain, in their own words, what a smart city is. We then provided a definition for all participants. We defined a smart city as "an urban area that uses different types of electronic methods and sensors to collect data. Insights gained from that data are used to manage assets, resources and services efficiently" [133]. We then asked participants to tell us whether they are aware of any cities with smart sensing technologies. We also asked participants to specify data collection scenarios in the city that would make them comfortable or uncomfortable.

(2) *Attitudes toward smart city data collection and use scenarios:* The goal of section two of the interview was to understand participants' attitudes and perceptions toward smart city data collection and use scenarios. To achieve this, we used the vignette methodology. Vignettes are "short stories about hypothetical characters in specified circumstances, to whose situation the interviewee is invited to respond," [49]. This approach has been commonly used in the literature to elicit participants' opinions in different contexts [41, 42, 88, 89]. Among the vignettes, we varied three factors that have been shown to influence people's concerns related to data collections: `data_type` [41, 124], `data_access` [10, 124], and `data_retention` [41]. The factors and their levels are presented in Table 1.
For each factor, we selected levels that we hypothesized to represent a wide range of concerns and perspectives. Based on the level of granularity, we selected two categories for `data_type`: high granularity and low granularity. We considered *real-time video footage from people* as the highly granular data. In addition, we included three low granular data types that are currently being collected by city-wide sensing technologies: detecting the presence of gunshots [86], collecting the number of people [111], and collecting air quality information [17]. While these low granular data types are considered to be aggregated or anonymized, they each have distinct implied sensing applications, namely crime, population, and environment, respectively, which relate back to city residents at varying degrees of association.

The second factor in the vignettes was `data_access`. We considered two categories of access: public access (open data) and restricted access. In the vignettes, we referred to *public access* by saying "everybody has access to the collected data." We considered two types of *restricted access*: public sector and private sector. For the public sector, we tested data being accessed by law enforcement officers, and by the mayor's office. We included data being accessed by the insurance companies for restricted access by the private sector.

Finally, we included `data_retention` in the scenarios. We considered two categories for retention time: short retention to be one day and long retention to be forever. It is important to note that we used the vignettes not to test specific use cases of smart city technologies and rather to facilitate in-depth conversations with participants to discuss each factor in the scenario more broadly. Therefore, we will not report on the findings related to each scenario, and instead, we will provide the most common themes among the scenarios.

In several cities, streetlamps are now equipped with sensors [17, 67]. In the presented vignettes, we used the streetlamp as the platform that sensors can be mounted on. Below is an example of a scenario that we presented to participants:

> *Imagine you are walking in the street where your home is located in Seattle, and you see a streetlamp. There are sensors on this streetlamp, which can only collect real-time video footage of people in the proximity of the streetlamp. Everybody can access this collected information. The collected information will never get deleted.*

Considering all the levels of all the factors, we can construct 32 vignettes (4 levels for `data_type`, 4 levels for `data_access`, and 2 levels for `data_retention`). To mitigate interview fatigue [107], each participant was randomly presented with a subset of 8 scenarios. The scenarios in each subset were selected in such a way that each participant would review all the levels of all the factors at least once. After presenting each data collection scenario, we asked participants questions to capture their level of concern, perceived benefits and harms, and desire to be notified about the data collection scenario.

3.2 Follow-Up Survey Study

Recruitment and compensation. We further enriched our understanding of participants' preferences and concerns by conducting an online follow-up survey. We conducted power analysis to identify the minimum required number of participants to be able to construct our planned statistical regression models. Based on the results of the power analysis, we recruited 348 participants from the Prolific crowdsourcing platform. We included participants who were living in the United States and had an approval rating of at least 95%. The survey took on average 14 minutes to be completed. We compensated each participant with US$5.

Survey procedure. By using the vignette methodology, we designed a within-subjects fractional-factorial survey to capture participants' concerns, attitudes, and expectations toward several smart city IoT data collection scenarios. We first provided participants with the consent form and asked them a few questions to obtain their consent to participate in our survey study. We then presented each participant with four randomly-assigned vignettes. We generated these scenarios similar to the interview study. We designed the scenarios by combining the levels of three factors of `data_type`, `data_access`, and `data_retention` (see Table 1). Each vignette described an IoT data collection and use scenario. Our survey has a fractional-factorial design. Out of 32 possible combinations, we randomly assigned each participant to a subset of scenarios, selected in such a way that each subset covers all the levels of all the factors. Since two of the factors had four levels, we needed to show at least four scenarios to cover all their levels. The reason we did not show more than four scenarios was to shorten the survey completion time and mitigate the survey fatigue [11] and keep the completion time under 15 minutes. We piloted our survey with five participants and realized that four scenarios is the maximum number of scenarios to keep the survey completion time below 15 minutes.

Similar to the interview study, in the survey study, we asked follow-up questions after presenting each scenario to obtain quantitative and qualitative findings. First, we asked participants to specify their level of concern about the data collection scenario on a 5-point scale, from not at all concerned to extremely concerned. We then asked open-ended questions to capture participants' reasons behind their level of concern.

We next asked participants to assess potential benefits of the presented data collection scenarios to themselves and to the society on a 5-point scale from not at all beneficial to extremely beneficial. We asked open-ended questions to better understand participants' justifications for their assessments. Similarly, we asked participants to assess how harmful the data collection and use vignettes could potentially be to themselves and to society.

As in the interview study, we asked participants to specify, on a 4-point scale, how often they would like to be notified about the described data collection scenario. Participants could select any of the four ordered frequencies: *i) every time this data is being collected about you, ii) every once in a while, when this data is being collected about you, iii) only the first time this data is being collected about you*, or *iv) never*. We then asked open-ended follow-up questions to better understand participants' preferences in receiving notifications about the IoT data collection scenarios. We ended the survey by asking demographic questions (e.g., age, gender, ethnicity, number of children in the household).

3.3 Data Analysis

We conducted qualitative data analysis to analyze participants' interview responses as well as responses to survey open-ended questions. We used statistical methods to quantitatively analyze participants' survey responses.

Qualitative analysis. All the interviews were audio recorded and transcribed by an IRB-approved third-party transcription service. We conducted the qualitative analysis on interview responses in two cycles [113]. In the first cycle, we used structural coding, which is particularly useful for coding responses from semi-structured interviews [85, 113]. We derived 5 structural main codes (e.g., smart city scenario assessment), which were divided into 23 sub-codes (e.g., harms), 21 sub-sub-codes (e.g., information misuse), and 3 sub-sub-sub-codes (e.g., discrimination). To further analyze the codes and find general themes and patterns, we applied thematic analysis [18]. Two researchers from the team had several hours of coding discussions to collaboratively create the codebook and find the overarching themes. After all the coding disagreements were resolved, and the coders agreed on a codebook, the primary coder applied the codebook to all the interview transcripts.

Since the interview questions subsumed the survey questions, we coded the survey open-ended responses using the codebook we constructed in the interview study. The codebook captured all the themes from the survey and, therefore, we did not add any new code to this codebook. We provide the final codebook in Appendix C.

Quantitative analysis. The survey study enabled us to further investigate participants' perceptions of smart city technologies by statistically modeling their concerns and preferences toward the presented IoT data collection vignettes.

Using cumulative link mixed models (CLMMs), we considered six dependent variables (DVs)—namely 1) concern about data collection, 2) benefit of data collection to self, 3) benefit of data collection to society, 4) harm of data collection to self, 5) harm of data collection to society, and 6) desired notification frequency about data collection—and built six regression models.

To build the models that best fit the data, we performed backward elimination [76]. In each stage of the model selection process, we compared the models by their Akaike information criterion (AIC), which is a general metric for the quality of the models in terms of goodness of fit [19]. We included all the independent variables (IVs), including the scenario factors (see Table 1) as well as the demographic factors (e.g., age, gender, education, income) and their two-way interaction terms in the first step of the backward elimination process. In addition, we included an ordinal categorical factor of `scenario_order` with four levels (e.g., first scenario) to control for the order in which the scenarios were presented to each participant. We gradually removed the factors that did not improve the model fit. Below is the union of the IVs that were included in the final best models:

- `data_type`: Nominal categorical factor to describe type of collected data with four levels of video footage, presence of gunshots, number of people, and air quality.
- `data_access`: Nominal categorical factor to describe who can access the collected data with four levels of everybody, law enforcement officers, mayor's office, and insurance companies.
- `data_retention`: Nominal categorical factor to describe the retention time with two levels of 1 day and forever.
- `income`: Ordinal categorical factor to describe the level of participants' annual income with 12 levels (e.g, $10,000 to $19,999). Since we did not have enough data in every income level in our collected dataset, to improve the statistical power of our models, we treated this factor as a numeric factor, where for each interval, we selected the mid-point and fed it to the model, after normalizing it by a base income value of $5,000.
- `# children`: Discrete numeric factor to describe how many children live in participants' homes.

Model baseline selection. The coefficients for each level (e.g., video footage) of categorical IVs (e.g., `data_type`) in the regression models should be interpreted compared to the *baseline* of that factor (e.g., air quality). Therefore, any level of factors can be selected as the baseline, and this decision does not have any impact on how the model fits the data. We selected the baseline for `data_type` to be air quality, the baseline for `data_access` to be law enforcement officers, and the baseline for `data_retention` to be 1 day.

In what follows, we provide details on the CLMM used to describe participants' level of concern with data collection scenarios. The construction of the remaining models (i.e., models to describe participants' assessments of data collection benefits and harms to self and society, and desired notification frequency) follow similar steps and is therefore omitted for brevity.

Consider the i^{th} observation with participant p_i, whose reported level of concern is denoted by a discrete random variable $Y_i \in \{1, \ldots, 5\}$. The probability that the concern level of the participant in this observation is at most $y \in \{1, \ldots, 4\}$ is modeled by the CLMM as

$$\Pr\left[Y_i \leq y\right] = \sigma\left(\alpha_{y|y+1} + \mu_{p_i} - \bar{\beta}_i\right), \tag{1}$$

where $\sigma(\cdot)$ denotes the sigmoid function, $\alpha_{y|y+1}$ denotes the threshold parameter between the two consecutive response levels y and $y+1$, and μ_{p_i} denotes the random effect corresponding to participant p_i, modeled as a Gaussian random variable with zero mean and variance σ_μ^2 determined by the model. Moreover, $\bar{\beta}_i$ is defined as

$$\begin{aligned}\bar{\beta}_i \coloneqq\ & \beta_{\texttt{income}} \cdot \texttt{income}_i \\ & + \beta_{\texttt{data\_type}_i} + \beta_{\texttt{data\_access}_i} + \beta_{\texttt{data\_retention}_i},\end{aligned}$$

where for a factor f, β_f denotes its corresponding model coefficient, and $\texttt{income}_i$, $\texttt{data\_type}_i$, $\texttt{data\_access}_i$, and $\texttt{data\_retention}_i$ denote the (numeric/categorical) levels of income, data type, access type, and retention time observed in the i^{th} observation, respectively.

Based on the cumulative concern level probabilities defined in (1), we denote the *odds ratio of being concerned* for a given categorical factor f with respect to its baseline f_{baseline} by $\text{OR}^f_{\text{concern}}$, and we define it as[2]

$$\text{OR}^f_{\text{concern}} \coloneqq \frac{\left(\frac{\Pr(\text{being concerned} \mid f)}{1-\Pr(\text{being concerned} \mid f)}\right)}{\left(\frac{\Pr(\text{being concerned} \mid f_{\text{baseline}})}{1-\Pr(\text{being concerned} \mid f_{\text{baseline}})}\right)}, \tag{2}$$

where $\Pr(\text{being concerned} \mid f)$ and $\Pr(\text{being concerned} \mid f_{\text{baseline}})$ denote the probabilities that a typical participant is concerned given the factor f and the factor baseline f_{baseline}, respectively. As we prove in Appendix D, the odds ratio in (2) can be written in closed form as

$$\text{OR}^f_{\text{concern}} = \exp\left(\beta_f\right). \tag{3}$$

3.4 Limitations

As is typical for qualitative research, due to the small number of interview participants, the findings of the interview study should not be generalized. We enhanced the generalizability of our study findings by conducting a large-scale follow-up survey and recruiting US participants on Prolific. The data from this platform has been

[2]For numeric factors, such as `income` and `# children`, we define the odds ratio as the ratio of the odds associated with one unit of increase in the factor.

Interviewee ID	Age	Gender	Race	Technical Background
I-P1	30	Female	Black or African American, White	No
I-P2	26	Female	Asian	Yes
I-P3	31	Female	White	No
I-P4	32	Female	Black or African American	No
I-P5	49	Male	Hispanic or Latino or Spanish Origin of any race	No
I-P6	28	Female	Asian, Native Hawaiian or Other Pacific Islander	No
I-P7	58	Female	White	Yes
I-P8	24	Non-Binary	Hispanic or Latino or Spanish Origin of any race, White	No
I-P9	39	Non-Binary	White	No
I-P10	71	Female	White	No
I-P11	61	Male	White	Yes
I-P12	43	Male	Hispanic or Latino or Spanish Origin of any race	Yes
I-P13	37	Female	Black or African American	No
I-P14	25	Male	Hispanic or Latino or Spanish Origin of any race	No
I-P15	64	Female	American Indian or Alaskan Native, White	No
I-P16	30	Female	Black or African American	Yes
I-P17	63	Male	White	Yes
I-P18	35	Female	Asian	No
I-P19	41	Female	White	No
I-P20	50	Female	Black or African American	Yes
I-P21	58	Male	White	Yes

Table 2: Overview of interview participants' demographic information. The complete demographic information of interviewees can be found in Table 8 in Appendix E.

Age		Gender		Race		Technical Background	
18-29	56%	Female	57%	American Indian or Alaskan Native	2%	Yes	27%
30-49	32%	Male	41%	Asian	11%	No	73%
50-64	10%	Non-binary	2%	Black or African American	9%		
65+	2%			Hispanic or Latino or Spanish Origin of any race	11%		
				Native Hawaiian or Other Pacific Islander	0%		
				White	74%		
				Other	1%		

Table 3: Overview of survey participants' demographic information. Note that the percentages under the "Race" column do not sum up to 100% as participants could choose any number of options. The complete demographic information of survey participants can be found in Table 9 in Appendix E.

shown to be more reliable than other crowdsourcing platforms, such as Amazon Mechanical Turk (MTurk) in terms of comprehension, attention, and honesty of its pool of participants [106, 109]. As a common limitation, crowdsourcing platforms are not completely representative of the general population. Nevertheless, they are commonly used in the literature to elicit participants' privacy attitudes and understanding [1, 44, 116].

In both interview and survey studies, we generated the IoT data collection scenarios based on three factors. Future studies should extend our work by exploring the impact of more factors (e.g., where the collected data is being stored, how the stored data is being protected) and more levels (e.g., the retention time of 1 month or 1 year) on participants' concerns and attitudes. Extending the number of generated vignettes would require collecting more participants to have enough statistical power for the quantitative analysis.

Interview participants are prone to biases [5]. As several of our participants acknowledged, the interview process helped them learn more about smart cities and their potential benefits and harms. Therefore, participants may have been more informed when assessing the benefits and harms in later scenarios as compared to the initial ones. We mitigated this potential order bias [16] in the survey study by randomly assigning participants to scenarios. In addition, in all regression models, we included the factor `scenario_order` to control for the order in which participants viewed the scenarios. Through the backward elimination process, this factor was removed from the final models as it did not have a significant impact on participants' attitudes and perceptions and did not help improve the model fit.

4 RESULTS

As part of our interview, we asked participants to specify what neighborhood they were living in. We excluded and did not analyze responses from two participants whose answers to this question did not match their screening surveys. We, therefore, report the findings from 21 participants. Table 2 shows a summary of, and Appendix E provides the complete interview participants' demographic information.

We then recruited 356 US participants from Prolific who had approval ratings of at least 95%. We removed eight participants who did not provide relevant answers to the open-ended questions and instead used the open-ended responses to advertise a product or service. The final dataset includes responses from 348 participants. We provide a summary of survey participants' demographic information in Table 3 and the complete demographic information of survey participants in Appendix E.

When providing quotes from participants, we refer to interview participants as I-P and survey participants as S-P. For example, I-P12 is interviewee number 12. When reporting the themes, we will provide representative quotes from both interview and survey participants.

4.1 Most Interview Participants Defined Smart Cities Around Technologies Rather than People

Despite their neighborhoods being publicly prioritized for smart city development by the city local government, most interview participants (14/21) reported that they had never heard of the term "smart cities." Nevertheless, we asked all participants to define, in their words, what a smart city means. When defining a smart city, interviewees who had not heard about smart cities prior to the interview, could not provide a concrete example of smart city technologies in the city under study. On the contrary, all the interviewees who had previously heard about the term smart city, provided examples of problem areas in the city or their neighborhood (e.g., potholes, crime) and explained how current or future smart city technologies could address those challenges.

Most participants (15/21) perceived a smart city as a technology-driven city with enhanced technological solutions in several domains, including transportation (e.g., autonomous cars), economy (e.g., cashless payments), environment (e.g., solar panels), and healthcare (e.g., telemedicine). Some participants (8/21) brought up internet connectivity in their definitions, and some (8/21) mentioned smart devices when describing smart cities, ranging from personal devices, such as smartphones or voice-activated smart home devices, to city-wide sensors, such as air quality sensors or security cameras.

There is a documented lack of attention to the people component in the narratives of smart cities in practice and in the literature [61, 66, 137]. Similarly, in our interview study, only a small number of participants (8/21) talked about the human aspect of smart cities by mentioning terms such as community, public, people, citizens, and society. For example, I-P21 reported:

> Compared to traditional cities, a smart city would have more community functions to get people together and to get people to interact with each other.

Some of the participants (3/8) who alluded to the human component of smart cities centered their definitions around accessibility and equity. I-P6 mentioned accessibility as the main component of smart cities:

> A city is smart where people who have difficulty with hearing or who have different visions are able to interact at the same level as people who don't think about those privileges.

4.2 Smart City Data Practices and Demographic Factors Influenced Participants' Attitudes and Concerns

We asked both interview and survey participants questions to understand their concerns, perceptions, and expectations toward smart city data collection vignettes. We found several similarities and important differences between the interview and survey participants in their assessment and attitudes concerning the presented smart city scenarios.

To analyze the survey responses, we conducted model selection using backward elimination and built cumulative link mixed models (CLMMs) that could best explain survey participants' concerns and expectations toward the smart city vignettes. In addition to the three scenario factors (see Table 1), we included participants' complete demographic factors in the models, including their age, gender, income, and education level. Although most demographic factors eventually got removed during the model selection process (see Section 3.3), the final models included a few demographic factors (e.g., level of income) that significantly influenced participants' smart city concerns and perceptions. We provide the complete regression results, where for each factor-level, we include the effect size (β), odds ratio (`OR`), standard error (`Std. Error`), and the p-`value`. The regression results for models describing participants' level of concern, assessment of potential benefits to self and society, assessment of potential harms to self and society, and the desire to receive notification about smart city scenarios are included in Tables 4, 5, 6, and 7, respectively. In all regression models, the type of collected data (`data_type`) was the most important factor, and the retention time (`data_retention`) was the least important factor in explaining survey participants' attitudes and concerns.

4.2.1 Participants Had More Privacy Concerns Toward the Collection of Data About People vs. Environment. Both interview and survey participants expressed higher levels of privacy concerns about smart city sensors collecting information about city occupants compared to environmental data being collected. Most interviewees (13/21) and almost all survey participants (275/348) found the collection of video footage in the proximity of city streetlamps to be the most privacy-invasive (see Table 4, row 1: estimate = 2.49, p-value < 0.001). Similarly, several interviewees (8/21) and survey (98/348) participants perceived the collection of data on the number of people in the proximity of the streetlamp to be privacy-invasive (see Table 4, row 2: estimate = 1.62, p-value < 0.001). Participants (interview (5/21) and survey (47/348)) were concerned about being tracked by this data collection. I-P5 mentioned:

> Collecting number of people is absolutely privacy invasive because even if it seems anonymous, knowing how many people are in a specific area is just a step away from more individual tracking and I do not want that in my neighborhood.

Some participants (interview (4/21) and survey (15/348)) wanted privacy-invasive smart city sensors, such as security cameras, to be easily visible to the public. S-P42 explained why they found video collection in a city to be extremely concerning, especially when not readily visible to the public:

> It's really concerning in the sense that it's capturing people's identifiable images without them really knowing/thinking about it. Are these cameras even visible on the streetlamps and can we easily spot them? Who knows how the footage is going to be

Row	Factor	Level of Concern (AIC = 3489)			
		OR	β	Std. Error	p-value
	data_type (baseline = Air quality)				
1	Video footage	12.02	2.49	0.18	***
2	Number of people	5.07	1.62	0.19	***
3	Presence of gunshots	3.00	1.10	0.17	***
	data_access (baseline = Law enforcement officers)				
4	Insurance companies	2.33	0.84	0.16	***
5	Everybody	1.76	0.56	0.16	***
6	Mayor's office	1.07	0.06	0.16	0.69
	data_retention (baseline = 1 day)				
7	Kept forever	1.39	0.33	0.11	**
	income_level (numeric)				
8	Income	0.96	−0.04	0.02	*
	Threshold Coefficients				
9	$\alpha_{1\|2}$	-	1.06	0.23	-
10	$\alpha_{2\|3}$	-	2.04	0.24	-
11	$\alpha_{3\|4}$	-	2.88	0.25	-
12	$\alpha_{4\|5}$	-	3.79	0.26	-
	Random Effects				
13	σ^2_μ	-	0.80	-	-

Note: $^{*}p < 0.05$ $^{**}p < 0.01$ $^{***}p < 0.001$

Table 4: CLMM regression model to describe how various factors impact survey participants' level of concern about smart city scenarios. Each row corresponds to a single factor and shows the resulting model estimate, i.e., coefficient, for that factor, alongside the odds ratio (defined in (2)), standard error (`Std. Error`), and p-value. The levels of factors (e.g., `data_type` in rows 1-3) are ranked in descending order according to their effect size, represented by the magnitude of the model coefficients (β). Note that a positive estimate for a factor-level (e.g., `data_type`:video footage) implies that transitioning from the baseline of the corresponding factor (e.g., air quality) to that level of the factor (e.g., video footage) would increase the perceived level of concern. A negative estimate reflects the opposite of this trend. We also include the AIC value for the model, which represents the model's goodness of fit.

> misused. This is kind of just what modern society is like these days.

On the contrary, collecting data about the environment as opposed to personal data posed almost no privacy concerns to our interview and survey participants. Among the four tested levels of `data_type`, the regression analysis showed that participants were least concerned about the collection of air quality data. The reason that most interview (12/21) and survey (187/348) participants provided when discussing their lack of privacy concern about air quality data collection was that this information was not perceived to be personally identifiable. I-P20 reported:

> I am really not worried about this information being collected in my city. No human data at all is collected, so I have no privacy concerns, and I don't believe it can be harmful to anyone.

Interview and survey participants preferred more restricted access to the data that are collected about people (video footage and the number of people), and reported to be more comfortable when such data are being shared only with *appropriate* parties as opposed to the public. S-P300 discussed why they were concerned about the data on the number of people being accessible by the public:

> If everyone could check it, what happens if a terrorist/criminal decides to set off something when there are the most people? They could check it each day.

Participants' assessment of which party should have access to the collected data was largely dependent on the trust they had in that party. S-P18 expressed their trust in law enforcement officers to have access to the collected video footage:

> I think law enforcement will do what they can to keep people safe so I believe the video footage will be in safe hands if shared with them.

Unlike data collected about people, participants preferred non-personal data about their environment to be accessible to the public. I-P1 provided justification as to why they preferred the air quality data to be accessible by everybody:

> This information can't be used to harm anyone, if anything it'll be more helpful for people who may have asthma or are sensitive to humidity.

Due to its privacy implications, permanent retention time became a significant concern when data about people were being collected in the smart city scenarios (see Table 4, row 7: estimate = 0.33, p-value < 0.01). However, when data about their environment were being collected, most interview (15/21) and survey (192/348) participants reported that they would prefer a longer retention time as opposed to one day of retention. S-P80 explained why they were concerned about the mayor's office having access to the video footage that will never get deleted:

> Never deleting data is very invasive of personal privacy. Data that is being kept forever could be potentially used in harmful ways, even if we do not see them now.

4.2.2 Participants Were Concerned About the Impact of Smart Cities on Marginalized Communities. Surveillance technologies, including city-wide security cameras, have been shown to pose greater harm to marginalized communities [52, 121]. Our study confirmed this unequal distribution of harm not only for surveillance technologies but also for other smart city technologies that have not been adequately discussed in prior research (e.g., air quality sensors).

Although only a subset of the presented smart city scenarios—primarily those involving the collection of data about people—posed privacy concerns to our interview and survey participants, they reported being concerned about the ethical implications of *all* the presented smart city data collections. All interviewees (21/21) and some survey participants (X/348) reported being concerned about the ethical implications of smart city scenarios for marginalized communities. Participants were concerned about the risk of potential discrimination and explained how their own neighborhoods or other underserved parts of the city could be harmed by the deployment of smart city technologies.

Among the presented smart city scenarios, participants were most concerned about the potential discrimination as a result of the collection of the presence of gunshots (interview (6/21) and survey (72/348)). Several participants (interview (6/21) and survey (47/348)) were concerned about the ethical implications of collecting air quality data in marginalized neighborhoods. Some participants (interview (5/21) and survey (44/348)) expressed concerns about the racial profiling caused by the collection of real-time video footage in communities of color, and a few (interview (3/21) and survey (35/348)) were concerned about the disproportionate harm of collecting the number of people in underserved neighborhoods.

S-P51 were concerned that lack of equity in the distribution of gunshot detectors could lead to increased discrimination against marginalized communities:

> I am actually concerned with gunshot collection. Well we should ask where is this technology is actually located? Is it located in an over-policed area that is mostly occupied by people of color? I think this would lead to more violence against already underprivileged demographics.

I-P9 expressed concerns about false positives of gunshot detectors in Black neighborhoods:

> Well, I don't live in a neighborhood that's affected by this. However, I would be concerned if a sound that wasn't a gunshot was interpreted to be a gunshot and the police were called, particularly if I was a black person because of problems with the police.

S-P112 discussed how collecting air quality data in underserved neighborhoods, including their own area, and sharing this information with insurance companies could be concerning:

> I would worry that it would affect the cost of insurance for people in my area neighborhood. I would worry that the most impoverished would be living in similar areas with the highest pollution, so an increase in insurance prices could cause a major issue for them.

Due to the potential discriminatory outcomes of such predictive policing technologies and to increase accountability and awareness, participants preferred this type of collected data to be shared with the public as opposed to specific entities. I-P2 discussed why they wanted the information on the presence of gunshots to be shared with the public:

> I believe information like this would be beneficial for the community as it would raise awareness and keep others informed of the situation.

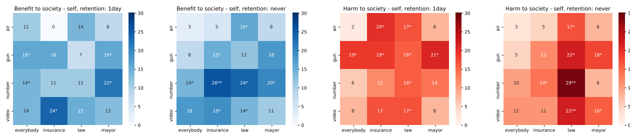

Figure 1: Percentage difference in perceived benefits and harms between society and self for various data collection and use scenarios. Statistically significant differences are marked with asterisks (*$p < 0.05$, **$p < 0.01$).

4.2.3 Differences in Models to Describe Perceived Harms and Perceived Benefits. We asked interview and survey participants questions to capture how harmful or beneficial they perceive the smart city scenarios to be for themselves and for society. Although some of the factors in the regression models to explain benefits and harms are the same (`data_type`, `data_access`), these regression models are not identical (see Tables 5 and 6). Regression results showed that `data_retention` is only effective in explaining survey participants' perceived harms (and not the benefits), while the `#Children` only impacts participants' perceived benefits (and not the harms) of smart city scenarios.

4.2.4 Perceived Impact of Smart Cities on Society is Different than on Individuals. Interview and survey participants perceived smart cities to have a larger impact on society than on themselves. Compared to its impact on themselves, interviewees provided more examples and spent more time discussing how that scenario could benefit/harm society. Similarly, survey participants perceived all 32 smart city data collection and use scenarios to be more harmful/beneficial to society compared to themselves (see Appendix F). For several scenarios, the difference between self and society was statistically significant (see Figure 1). For example, participants found almost all scenarios involving the collection of the presence of gunshots to be more harmful to society than to themselves.

4.2.5 Participants Perceived a Trade-off Between Privacy and Safety in Smart Cities. Our qualitative analysis showed that several participants viewed preserving their personal privacy and having safe neighborhoods as a trade-off in smart cities. When discussing the benefits and harms of video footage collection, several interviewees (7/21) and survey participants (107/348) found such data collection to be beneficial to society as it improves the safety of their neighborhoods. Most participants who mentioned safety as the benefit of video collection, reported having privacy concerns with such data collection. In fact, interview and survey participants found the collection of video footage to be most harmful to self (see Table 6, row 1: estimate = 2.56, p-value < 0.001). I-P17 reported that despite having concerns about video collection, they still found it to be potentially beneficial:

> I don't agree with this surveillance scenario. That said they still are an effective tool for investigating crimes.

Row	Factor	Beneficial to Self (AIC = 3428)				Beneficial to Society (AIC = 3641)			
		OR	β	Std. Error	p-value	OR	β	Std. Error	p-value
	data_type (baseline = Air quality)								
1	Number of people	0.12	−2.15	0.18	***	0.11	−2.16	0.18	***
2	Video footage	0.22	−1.52	0.16	***	0.22	−1.53	0.16	***
3	Presence of gunshots	0.56	−0.59	0.15	***	0.65	−0.43	0.15	**
	data_access (baseline = Law enforcement officers)								
4	Insurance companies	0.30	−1.22	0.17	***	0.21	−1.55	0.17	***
5	Mayor's office	0.68	−0.39	0.15	*	0.56	−0.57	0.15	***
6	Everybody	0.96	−0.04	0.16	0.82	0.72	−0.33	0.16	*
	#children (numeric)								
7	#Children	1.35	0.30	0.09	***	1.26	0.23	0.09	*
	Threshold Coefficients								
8	$\alpha_{1\|2}$	-	−1.26	0.30	-	-	−2.59	0.33	-
9	$\alpha_{2\|3}$	-	−0.05	0.30	-	-	−1.29	0.32	-
10	$\alpha_{3\|4}$	-	0.99	0.30	-	-	−0.05	0.32	-
11	$\alpha_{4\|5}$	-	2.16	0.31	-	-	1.23	0.32	-
	Random Effects								
12	σ^2_μ	-	0.76	-	-	-	1.14	-	-

Note: $^{*}p < 0.05$ $^{**}p < 0.01$ $^{***}p < 0.001$

Table 5: CLMM regression models to describe how various factors impact survey participants' assessments of benefits of smart city scenarios to self and society. Each row corresponds to a single factor and shows the resulting model estimates, i.e., coefficients, for that factor, alongside the odds ratios (defined in (2)), standard errors (`Std. Error`), and p-values. The levels of factors (e.g., `data_type` in rows 1-3) are ranked in descending order according to their effect size, represented by the magnitude of the model coefficients (β). Note that a negative estimate for a factor-level (e.g., `data_type`:video footage) implies that transitioning from the baseline of the corresponding factor (e.g., air quality) to that level of the factor (e.g., video footage) would decrease the perceived benefit of that data collection. A positive estimate reflects the opposite of this trend. We also include the AIC values for the models, which represent the models' goodness of fit.

Row	Factor	Harmful to Self (AIC = 2605)				Harmful to Society (AIC = 3111)			
		OR	β	Std. Error	p-value	OR	β	Std. Error	p-value
	data_type (baseline = Air quality)								
1	Video footage	12.94	2.56	0.22	***	12.16	2.50	0.20	***
2	Number of people	5.31	1.67	0.23	***	4.80	1.57	0.21	***
3	Presence of gunshots	2.54	0.93	0.22	***	3.02	1.10	0.19	***
	data_access (baseline = Law enforcement officers)								
4	Insurance companies	3.35	1.21	0.20	***	2.06	0.72	0.18	***
5	Everybody	3.00	1.10	0.20	***	1.62	0.48	0.18	**
6	Mayor's office	1.41	0.34	0.20	0.09	0.93	−0.07	0.17	0.68
	data_retention (baseline = 1 day)								
7	Kept forever	1.33	0.29	0.13	*	1.45	0.37	0.12	**
	Threshold Coefficients								
8	$\alpha_{1\|2}$	-	2.90	0.26	-	-	1.71	0.22	-
9	$\alpha_{2\|3}$	-	3.94	0.28	-	-	3.03	0.24	-
10	$\alpha_{3\|4}$	-	4.97	0.30	-	-	4.04	0.26	-
11	$\alpha_{4\|5}$	-	5.82	0.32	-	-	4.92	0.28	-
	Random Effects								
12	σ^2_μ	-	1.23	-	-	-	1.66	-	-

Note: $^{*}p < 0.05$ $^{**}p < 0.01$ $^{***}p < 0.001$

Table 6: CLMM regression models to describe how various factors impact survey participants' assessments of harms of smart city scenarios to self and society. Each row corresponds to a single factor and shows the resulting model estimates, i.e., coefficients, for that factor, alongside the odds ratios (defined in (2)), standard errors (`Std. Error`), and p-values. The levels of factors (e.g., `data_type` in rows 1-3) are ranked in descending order according to their effect size, represented by the magnitude of the model coefficients (β). Note that a positive estimate for a factor-level (e.g., `data_type`:video footage) implies that transitioning from the baseline of the corresponding factor (e.g., air quality) to that level of the factor (e.g., video footage) would increase the perceived harm of that data collection. A negative estimate reflects the opposite of this trend. We also include the AIC values for the models, which represent the models' goodness of fit.

> I still think there's a bunch of problems when it comes to video surveillance data, but despite my concerns it's hard to argue with solved crimes.

4.2.6 Existence of Smart City Technologies Implies Underlying Challenges. Regardless of their data practices, some survey participants (21/348) found the very existence of smart city technologies in their neighborhoods to be harmful. Although all interview participants (21/21) acknowledged the needs of their neighborhoods when assessing the presented smart city vignettes, some of the survey respondents did not see challenges in their neighborhoods that need to be fixed by smart city technologies. These participants reported that such technologies being deployed in their neighborhoods could imply that the neighborhood has challenges (e.g., lack of safety, increased gun violence) that should get fixed. They found this particularly harmful as it could impact adversely impact people's perception of their neighborhoods, which could in turn lead to financial and economic losses. S-32 discussed how their neighborhood, which they perceived to be safe, could be harmed by the presence of gunshot detectors:

> We are living in a good and safe neighborhood, so there is no need for this type of tech here. I think by putting gunshot detectors around the neighborhood people would become more afraid of certain areas, and they may even decide to leave and work elsewhere.

4.2.7 Participants Found Raising Public Awareness as a Benefit of Open Data in Smart Cities. Some interview and survey participants mentioned capturing evidence in smart cities as a potential benefit of such deployments. For example, interview (7/21) and survey (81/348) participants perceived having smart city data accessible by the public as the second most beneficial level of `data_access`, mainly due to increased public awareness (see Table 5, row 6: beneficial to self: estimate = -0.04, p-value = 0.82, beneficial to society: estimate = -0.33, $p\text{-value} < 0.05$). I-P5 specified informing the public about the homelessness problems in their neighborhood as a positive outcome of everybody having access to the collected real-time video footage:

> I think it would be a good thing to wake up a lot of people who are in denial about what's going on here and have watched the insanity, like a zoo, that's in front of the courthouse 24/7. I think that could be a good thing and the more you can pull those folks out of denial, the better.

4.2.8 Participants Specified Information Overload and Exposure as a Harm of Open Data in Smart Cities. For interview and survey participants, the most frequently mentioned harm of everybody having access to the collected data was information overload (interview: 13/21 and survey: 76/348), which could lead to increased anxiety among people. I-P21 talked about information overload when the presented scenario mentioned that air quality data are shared with everybody:

> What do you do with that information? I mean, we know what the air quality is, it's not going to change. So, really what good is the information other than to make people upset. 'Oh, look at our air quality, it's horrible. Oh, this is bad.' Then, people get in an uproar, and really there's nothing we could do about it. You want to change the air quality, you got to make different steps to change air quality around, right?

Another commonly mentioned harm among our interview (8/21) and survey (27/348) participants was the risk of city residents being exposed to sensitive information as a result of open data in smart cites. Participants were particularly concerned about potential psychological harm of real-time video footage being easily accessible to everyone. S-18 discussed how exposing the public to crime-reletaed video footage could be harmful:

> I think there are still a lot of ethical questions about making this data public that has yet to be answered. Let's say there is an officer-involved shooting in an area covered by cameras. Police will likely have a press conference that releases footage. A likely similar situation for any other newsworthy crime. What right do we have to make the public see someone die? Could that be harmful? Could that increase the violence itself? I think to some extent, yes.

4.2.9 Participants Found Both Short And Long Retention Time to be Potentially Harmful. When discussing the impact of retention time on the benefits and harms of smart cities, our participants frequently mentioned that data should be retained as long as needed, and they perceived harms when this policy was violated. Such harms included privacy violations when collecting data for too long (see Table 6, row 7: harmful to self: estimate = 0.29, $p\text{-value} < 0.05$, harmful to society: estimate = 0.37, $p\text{-value} < 0.01$) or wasting resources (e.g., time, money) when retaining data for a short period that does not satisfy the purpose of data collection (interview (6/21) and survey (39/348)). S-P118 explained why they do not see any benefits in retaining the collected air quality data for 1 day:

> It seems like a waste of resources to collect air quality data that can only be seen by one person for only 24 hours. Are tax payers going to pay for such limited technology?

Similarly, I-P17 mentioned waste of time as a potential harm of the collected video footage being kept for 1 day:

> I think then it's a waste of time. I really do, because you want to be able to go back and look at things. I think that's the whole reason you do it, right? Somebody's going to sit there and watch it the whole time? No. You don't need that. You need it recorded so if something happens, we can go back to that time and date and look at it.

4.2.10 Participants With More Children Perceived More Benefits for Smart Cities. No demographic information was effective in the models to explain survey participants' perceived harms of smart city scenarios to themselves and to society (see Table 6). However, we found the number of children under the age of 18 to impact participants' benefit assessment. The regression coefficients showed that those living with more children under the age of 18 found IoT data collection scenarios to be significantly more beneficial to themselves (see Table 5, row 7: beneficial to self: estimate = 0.30, $p\text{-value} < 0.001$, beneficial to society: estimate = 0.23, $p\text{-value} < 0.05$) compared to those living with fewer children. Several participants mentioned their concerns about their children's safety and health and how IoT technologies (e.g., video cameras, air quality sensors) in their neighborhoods could mitigate those concerns.

Row	Factor	Notification Frequency (AIC = 2931)			
		OR	β	Std. Error	p-value
	data_type (baseline = Air quality)				
1	Video footage	4.31	1.46	0.18	***
2	Presence of gunshots	2.48	0.91	0.17	***
3	Number of people	1.94	0.67	0.19	***
	data_access (baseline = Law enforcement officers)				
4	Everybody	1.57	0.45	0.18	*
5	Insurance companies	1.37	0.32	0.18	*
6	Mayor's office	0.77	−0.26	0.17	0.13
	income_level (numeric)				
7	Income	0.93	−0.07	0.03	**
	Threshold Coefficients				
8	$\alpha_{1\|2}$	-	−0.93	0.26	-
9	$\alpha_{2\|3}$	-	−0.32	0.26	-
10	$\alpha_{3\|4}$	-	−0.88	0.26	-
	Random Effects				
11	σ^2_μ	-	2.27	-	-

Note: $^{*}p < 0.05$ $^{**}p < 0.01$ $^{***}p < 0.001$

Table 7: CLMM regression model to describe how various factors impact the frequency that survey participants would like to be notified about smart city scenarios. Each row corresponds to a single factor and shows the resulting model estimate, i.e., coefficient, for that factor, alongside the odds ratio (defined in (2)), standard error (Std. Error), and p-value. The levels of factors (e.g., data_type in rows 1-3) are ranked in descending order according to their effect size, represented by the magnitude of the model coefficients (β). Note that a positive estimate for a factorlevel (e.g., data_type:video footage) implies that transitioning from the baseline of the corresponding factor (e.g., air quality) to that level of the factor (e.g., video footage) would increase the desired frequency of receiving notification about that data collection. A negative estimate reflects the opposite of this trend. We also include the AIC value for the model, which represents the model's goodness of fit.

4.2.11 Participants Wanted Transparency and Autonomy in Smart Cities. We asked participants to specify their level of interest in being informed about smart city scenarios. We also asked interview and survey participants to discuss what aspects of the presented smart city scenarios they would like to be informed about if any. We surfaced three main categories of information that interview and survey participants were interested in being informed about, namely 1) privacy and data practices, 2) the impact of scenarios on themselves and society, and 3) the availability of user controls.

Almost all interview (16/21) and most survey (197/348) participants reported that they would like to get notifications about the *privacy and data practices* of smart city technologies in their neighborhoods. The desired privacy factors our participants mentioned in their open-ended responses were type of collected data (interview (16/16) and survey (166/197)), who the collected data will be shared with (interview (14/16) and survey (141/197)), the purpose of data collection (interview (14/16) and survey (130/197)), how often the data collection occurs (interview (9/16) and survey (92/197)), and where the smart city technology is located at (interview (8/16) and survey (43/197)). A few interview (3/16) and survey (17/197) participants expressed interest in knowing about data retention time. Indeed, this factor did not have a significant impact on survey participants' desired frequency of receiving notifications about smart city scenarios (see Table 7).

Several participants (interview (15/21) and survey (172/348)) reported that they would like to know about the potential benefits and harms of smart city scenarios to themselves and their neighborhoods. In addition, participants wanted to know about which city stakeholders are responsible for mitigating the harms of such technologies (interview (10/15) and survey (76/172)), and what concrete actions they are taking toward that (interview (9/15) and survey (66/172)).

Moreover, some interview (11/21) and survey (79/348) participants reported that they would like to have autonomy in smart cities and wanted to know about what *controls* they have, if any, over the smart city technologies. The desired types of controls our participants explicitly mentioned in their responses were being able to *opt-in or out of* the data collection (interview (10/11) and survey (65/79)) and data sharing (interview (10/11) and survey (61/79)), ability to *view* (interview (8/11) and survey (43/79)), *correct* (interview (7/11) and survey (38/79)), and *delete* (interview (7/11) and survey (31/79)) the collected data, and the ability to *provide feedback* (interview (6/11) and survey (30/79)) about the smart city scenarios. In addition, if any control was being provided, participants wanted to be informed about *what steps* (interview (10/11) and survey (74/79)) they need to take to exercise their autonomy.

A few interview (4/21) and survey (18/348) participants reported that even if they are not able to stop the city data collections from happening, having information about the location and the privacy and data practices of smart city technologies could potentially enable them to avoid being captured by such technologies. S-P45 talked about having a sense of control when being informed about city security cameras:

> I would want signs, I would want auditory warnings, I would want street lamps labeled with their privacy information. I want as much as possible done so that I can at least have somewhat of a semblance that I am choosing/not choosing to be recorded.

4.2.12 Participants Wanted More Frequent Notifications About Privacy-Concerning Smart City Scenarios. Having concerns about the privacy implications of smart city scenarios was the most frequently mentioned reason as to why participants (interview (13/21) and survey (87/348)) were interested in receiving notifications about those scenarios. The quantitative analysis indicated that survey participants' level of concern with smart city scenarios had a statistically significant and positive correlation (p- value < 0.001) with how often they would like to be informed about such smart city vignettes. Interview and survey participants were least interested in being notified about the collection of air quality data and most interested in getting notifications about the collection of their video footage in the city. S-P18 explained why they did not want to be notified about the collection of air quality data:

> Because this data collection really doesn't concern me or my privacy. I would like this kind of information to be made available to the members of the public, if they ever want to view it, but that doesn't mean I have to be notified about it.

Similarly, I-P11 expressed having privacy concerns with their collected video footage being shared with the public, and explained why they wanted to be notified every time such data were being collected:

> I would want to be aware of when my actions are being recorded and potentially watched by other people.

4.2.13 Experience and Awareness Influenced Smart City Assessments. We recruited our interview participants from the most underserved neighborhoods with the highest priority for smart city project deployments. Among several other factors, income and race were two indicators that the city council used to specify these underserved neighborhoods. Therefore, the neighborhoods that we recruited from were primarily occupied by low-income people of color.

Our quantitative and qualitative analysis showed that interview and survey participants' ability to relate to the smart city scenarios had a strong impact on their perceived concerns and preferences toward such data collection vignettes. When assessing the presented scenarios, interviewees explained how the technology could improve or worsen the challenges (e.g., high level of air pollutants, gun violence, lack of safety) of their neighborhoods, and they were all able to mention at least one benefit and one harm for each of the smart city scenarios. Indeed, their living experience resulted in their ability to relate to the presented smart city vignettes.

Survey respondents were less homogeneous in how they related to the presented data collection scenarios. Some participants (31/348) referred to the challenges of their own neighborhoods when specifying the potential benefits and harms of the smart city vignettes. Several participants (62/348) related to the scenarios by discussing how the presented scenarios could potentially benefit or harm other neighborhoods, mainly low-income communities of color. S-P27 discussed how insurance companies having exclusive access to air quality data could adversely impact low-income individuals:

> This situation would be very concerning and harmful to lower-income people because they can't simply manage to move to places where the air quality is better to get a better insurance rate.

Some survey respondents (49/348) were not able to identify any benefit or harm for some of the smart city scenarios. The majority of these participants, in their open-ended responses, indicated that their neighborhoods are not facing the challenges that the scenarios are aiming to help with. Therefore, they would perceive those technologies as a waste of resources if being deployed in their neighborhoods.

Confirming the qualitative findings, the regression analysis showed that survey participants' level of income significantly influences their attitudes and assessments toward smart city scenarios. In particular, lower-income participants expressed a significantly higher concern toward data collections compared to higher-income participants (see Table 4, row 8: estimate = -0.04, p-value < 0.05). We further broke down the concern into two main categories of concern over privacy and concern over the ethical implications of the scenario. Next, we explored the impact of income on each of these two categories of concern. By conducting the Kruskal-Wallis nonparametric test of correlation, we found that level of income only influences ethical concerns (p-value < 0.05) and has no significant impact on participants' concerns over privacy. This finding is in line with our qualitative results, where we highlighted that interview participants, who were all from low-income neighborhoods, reported being concerned about the ethical implications of presented scenarios, even when perceiving little or no privacy concerns. Moreover, the regression analysis showed that lower-income survey participants had a significantly stronger desire to receive notifications about smart city technologies compared to higher-income participants (see Table 7, row 7: estimate = -0.07, p-value < 0.01). This could be due to having greater perceived concerns and, therefore, a stronger desire to know more about the concerning smart city scenarios.

5 DISCUSSION

By conducting the interview and survey studies, we captured participants' nuanced understanding, concerns, and preferences toward smart city technologies. We start this section by discussing the factors that could potentially influence people's acceptability of smart city technologies. We then provide recommendations on designing human-centered and privacy-respecting smart cities. When distilling recommendations, we do not mention a specific stakeholder group because different cities might delegate responsibilities differently. Instead, we encourage smart city designers to consider our recommendations deeply and evaluate them in the context of their cities and the stakeholders therein.

5.1 Barriers and Incentives in Smart City Technology Acceptability

People's acceptability of smart city projects has proved to be a strong indicator of success or failure for several high-profile smart city initiatives, including the Sidewalk Toronto project [26, 63, 75], smart street lights in San Diego [47, 118], and the Replica mobility project in Portland [126]. Although our interview and survey studies were not designed to quantitatively measure people's smart city technology acceptability, participants' open-ended responses revealed several factors that could influence the public's acceptability of city-wide data collection and use scenarios.

Data privacy concerns. Our interview and survey participants reported having privacy concerns when data about people (video footage and the number of people) were being collected in the presented smart city scenarios (see Section 4.2.1). Participants were primarily concerned about their sensitive and private information (e.g., location) being disclosed by such data collections. Due to such concerns, participants preferred the collected data about people to only be shared with appropriate parties (e.g., law enforcement officers) and not the public at large. In addition, compared to other data types, participants were most interested in receiving notifications about collections of data types they perceived to be concerning.

Collecting people's data in urban environments is in conflict with people's right to privacy [59, 131], which could lead to people's rejection of surveillance technologies in public locations [8]. Acknowledging the privacy and safety trade-off [3, 9, 28], many of our interview and survey participants found the smart city technologies that collect people's data (video footage and the number of people) to be privacy-invasive. Still, they mentioned the potential benefits of these technologies to enhance the safety of their neighborhoods (see Section 4.2.5).

Social equity concerns. The collected data being potentially used for discriminatory purposes was another common concern our interview and survey participants reported having when assessing the presented smart city data collection and use vignettes (see Section 4.2.2). Our participants were particularly concerned about the potential discrimination that is caused when smart city technologies are heavily deployed in neighborhoods with a larger population of Black or African American residents. Participants' fear of discrimination was most frequently mentioned in scenarios involving predictive policing technologies (e.g., gunshot detectors). This finding speaks to the well-documented racial bias and police brutality that has led to a lack of trust in police forces among communities of color [71, 114]. To increase accountability in data handling and mitigate the potential harm to marginalized communities, participants wanted such data collections to be accessible to the public as opposed to being controlled by specific entities.

People's perception of their neighborhoods. Prior research has shown that people's perceived level of safety in their cities and neighborhoods strongly influences their acceptability of crime surveillance technologies [123]. We also found that our interview and survey participants' perception of their neighborhoods was a key factor impacting how they assessed the benefits, harms, and risks of smart city data collection and use scenarios. Open-ended responses suggested that participants' perceptions of their neighborhoods were shaped by either experiencing the challenges (e.g., lack of safety, high levels of air pollutants, gun violence) that the smart city technology (e.g., security cameras, air quality sensors, gunshot detectors) is aimed to solve, or having an awareness of those challenges without personally experiencing them. While all of our interview participants were able to relate to the presented smart city scenarios through having a personal experience, survey responses indicated varying levels of experience and awareness. Some survey participants reported personally experiencing the city challenges, several were aware of such challenges and their impact on marginalized communities and neighborhoods, and some were not able to relate to the presented smart city scenarios either through personal experience or awareness. Those participants who perceived the challenges in their neighborhoods and were able to relate to smart city scenarios provided examples of the potential benefits of smart city technologies. On the contrary, survey participants who reported perceiving no benefits in such technologies reported that deploying such technologies would be a waste of "tax payers' money."

Trust in technology and data practices. People's level of trust in technology and its privacy practices has been shown to significantly influence their technology acceptability and adoption [20, 21, 36, 57]. Our interview and survey participants expressed varying levels of trust toward the presented technologies and their data practices. The smart city technology that caused the most distrust among participants was the gunshot detectors. Several participants, the majority of whom were Black or African American, questioned the accuracy of these technologies [33, 58, 117] and how they could potentially be used to discriminate against communities of color (see Section 4.2.2).

Trust in technologies' data practices was another factor influencing participants' concerns and attitudes toward presented smart city scenarios. Several of our interview and survey participants did not trust the mayor's office to have access to any data that are collected by smart city technologies (see Section 4.2.1). Such distrust in government stakeholders has been shown to strongly influence people's acceptability of digital surveillance technologies during the COVID-19 pandemic [73, 130]. Another entity that most of our interview and survey participants did not trust was insurance companies. Our participants reported that they do not trust private organizations as they do not have the public's best interest in mind when collecting and processing their data. We also observed a divide among participants when assessing the benefits and harms of data access by law enforcement officers. Although many participants were comfortable with law enforcement officers having access to safety-related data (e.g., real-time video footage), several participants, primarily non-white participants, expressed concerns related to law enforcement having access to such information and potentially misusing this data against their communities (see Section 4.2.1).

Psychological burden. When assessing the harms of smart city scenarios, interview and survey participants mentioned some of the psychological challenges that city technologies could impose. To pave the path toward smart city adoption, such challenges and their harms should be carefully investigated and mitigated. The most frequently mentioned psychological burden was information overload as a result of open data in smart city projects. Although participants acknowledged the enhanced accountability of sharing the collected data with the public, several participants were concerned about the information overload that could lead to anxiety and mental fatigue (see Section 4.2.8). In addition to information overload, participants were concerned about the harm of public exposure to sensitive data collected by smart city technologies, such as raw video footage that could contain acts of violence. Some participants reported that making this information publicly available could lead to further crime and violence in their cities (see Section 4.2.8).

Prior research has identified the fear of being under "general suspicion" as one of the ethical implications of crime surveillance technologies [91], which could then become a roadblock in the acceptability of such technologies in cities [123]. Similarly, students perceived a significant distrust toward the deployment of surveillance security cameras in schools [15]. Our interviewees and survey participants referred to such fear when assessing the harms of security cameras and gunshot detectors being deployed in their neighborhoods. Participants reported that even the existence of these technologies in neighborhoods could imply a general distrust that the area is not safe and that it needs to be monitored. Some participants reported that such negative perceptions could particularly harm underserved neighborhoods by deterring future home seekers and businesses.

5.2 City Stakeholders Can Do More to Prioritize Privacy and Ethics in People-Centric Smart Cities

Others have argued that, for too long, smart cities have been a *surveillance theater*, where the focus of city stakeholders is more on vast data collections that enable innovation and technological advances and less on their desirability and impact on the city residents [95]. Changing such an economically and profit-driven narrative requires city stakeholders to start prioritizing people in smart cities over the smartness of cities. Based on our quantitative and qualitative findings, we provide actionable recommendations for city stakeholders to consider when designing people-centric smart city projects.

Providing equitable transparency and autonomy in smart cities. Interview and survey participants expressed interest in being informed about three aspects of smart city scenarios (see Section 4.2.11), namely: 1) privacy and data practices (e.g., type of collected data, purpose of data collection), 2) the impact of scenarios on themselves and the society (e.g., potential benefits and harms of scenarios), and 3) the availability of controls (e.g., opt-in/out of smart city data collection).

We found that participants' level of interest in receiving notifications depended largely on their level of privacy concerns toward such technologies (see Section 4.2.12). Participants desired to receive more frequent notifications about smart city technologies they perceived to be more concerning and vice versa. This finding suggests that the mode of transparency should respect nuances in people's privacy preferences and empower them to specify what scenarios they would like to be notified about. A Privacy Assistant [23, 30, 84, 112, 127] can be designed to enable such configurable information communication based on users' privacy preferences. These technologies can notify people about the security and privacy practices of nearby IoT technologies they are most concerned about without overwhelming them. Designing such technologies requires an in-depth knowledge of people's concerns and preferences toward smart city data collections. Our qualitative data and statistical models provide a foundation for the knowledge required to develop effective Privacy Assistants.

Moreover, our quantitative analysis showed that participants' socio-economic status has a significant impact on their desire to get notifications about smart city scenarios (see Section 4.2.13). We found that those with higher income levels prefer to receive less frequent notifications compared to lower-income participants. Designing equitable and inclusive information communication tools requires city stakeholders to carefully study and consider differences in the demographics of city residents.

Equitable distribution and visible smart city technologies. Our participants expressed concerns about (see Sections 4.2.1 and 4.2.2) and wanted to have transparency over (see Section 4.2.11) the location of smart city technologies. Open-ended responses surfaced two categories of location-related concerns, namely: 1) policies around technology distribution (see Section 4.2.2), and 2) visibility of technologies (see Section 4.2.1).

Participants were concerned about the distribution policies of technologies that capture information related to safety and crime, such as gunshot detectors (see Section 4.2.2). Respondents reported that the large number of such technologies in neighborhoods populated by marginalized communities could perpetuate disproportionate harms (e.g., heightened police presence) to those communities that are being captured by these smart technologies. To mitigate this, city planners and policymakers who are in charge of deciding the locations of sensors should listen to communities' concerns and take that into consideration to ensure that their sensor distribution policy will not cause harm to marginalized communities and exacerbate the already concerning and rising discrimination. Currently, 19 cities across the United States, including San Francisco, Seattle, and San Diego, have passed Community Control Over Police Surveillance (CCOPS) laws to empower the people in the city to decide if and how surveillance technologies should be used in their cities [6].

In addition, participants expressed concerns about the visibility of privacy-sensitive smart city technologies (see Section 4.2.1). To mitigate such concerns and raise public awareness, city stakeholders should make the recording sensors more visible and easily detectable. For example, Array of Things (AoT) [102], a smart city project in Chicago, uses elements based on designs from the School of the Art Institute of Chicago so as to capture public's attention. Moreover, making smart city technologies visible could lead to enhanced acceptability of such technologies as prior research has shown that surveillance technologies that are visible in public locations are more acceptable by people compared to invisible ones [123].

The accessibility of sensor indicators should be an important consideration for designers of these technologies. For example, in order to inform blind or low-vision individuals of smart city sensors, these technologies (e.g., security cameras) could use audible signals when the recording is in progress, similar to beeping crosswalks.

Making smart cities about people, not technologies. When asking interview participants to define smart cities, the majority of definitions were around the technology aspect of smart cities. Only a few participants mentioned the human element of smart cities in their definitions, such as accessibility and equity (see Section 4.1). Indeed, the rhetoric around smart cities, either from companies or the academic literature [7, 128], has missed people's concerns, perspectives, and attitudes and instead has primarily focused on the technological advances and resulting benefits of IoT-enabled cities. Our findings indicated that participants were significantly concerned about the privacy and ethical implications of smart cities (see Sections 4.2.2 and 4.2.1) and perceived several harms with the deployment of smart city technologies.

Although prior research has identified privacy and ethical implications as priority considerations for future smart cities [39, 124], to date, city stakeholders have primarily viewed people's privacy and ethical concerns as afterthoughts, and certainly not from the early stages of smart city planning and deployment [37]. Moreover, prior research has frequently highlighted the importance of bottom-up and participatory smart city designs [34, 35, 38, 54, 55, 68, 69, 94, 132]. However, none has focused on surfacing privacy and ethical concerns to be considered in such desired co-creation process. To ensure designing privacy-respecting and equitable smart cities that work for and are welcomed by their residents, it is critical for city stakeholders to actively seek all people's input and hear their concerns in all the stages of smart city projects, from brainstorming to deployment.

Moving away from a narrow focus. Due to the extensive data collection and processing, smart cities are essentially surveillance cities [99]. The academic literature and conversations around surveillance and smart cities have been primarily occupied by narrow scenarios and single technologies [78], mostly involving security cameras. Although surveillance cameras pose undeniably huge privacy and ethical risks, fixating on a single technology and scenario can be of far greater harm.

When we asked our interview participants to define smart cities and provide examples of smart city technologies, several participants were only able to mention security cameras, even though several other technologies existed in their neighborhoods. One could hypothesize that such limited awareness could be attributed to people's significant concern toward security cameras compared to other city technologies. However, our interview and survey studies proved that this hypothesis is not entirely correct. When presenting participants with diverse smart city scenarios involving various IoT technologies and various parties accessing the collected data, participants started to critically assess the scenarios' potential risks and benefits, no less than scenarios involving security cameras.

On the other hand, several interview and survey participants reported having no concerns about city-wide security cameras as they have been "used to" such technologies and, therefore, do not see them as risky anymore. Prior research has referred to this behavior and mindset as habituation [110], which could significantly reduce people's perceived risks of technologies [45, 51].

For communities to effectively participate in the smart city decision-making process, they need to be empowered to criticize the smart city technologies, and that requires perceiving the risks of such technologies. Therefore, city stakeholders need to bring the traditional and forgotten smart technologies back to the table and incentivize communities to openly discuss and criticize them [134]. It is only by doing so that city residents can truly understand the potential harms or values of such project proposals and be able to make an informed contribution to the future of their cities.

Not losing individuals for the society. Our participants were more informed about and better able to discuss the impact that smart cities could have on the society compared to their potential impact on themselves (see Section 4.2.4). For several scenarios, such perceived difference between societal and individual impacts were statistically significant (see Figure 1). Although the perceived discrepancy could potentially lead to technological advances due to individual sacrifices, this skewed view, if not informed, could be abused to influence the adoption of smart city technologies.

To prevent pro-social narratives from becoming a Trojan horse for individuals' rights and preferences sacrifices, city stakeholders should openly discuss and be transparent about how smart city technologies influence individuals, and what harms and benefits people should expect to see on a personal level, in addition to the societal level. This could be especially valuable to disclose when the impacts on individuals and the society are not consistent and there are trade-offs and tensions between them. People should be informed about such tensions, how city stakeholders handle the tensions, and what controls they are being provided with to manage their concerns.

6 CONCLUSION

A growing number of cities around the world are designing and deploying technologies in their cities to address the monumental challenges of urbanization. The ubiquitous smart city technologies are fueled by collecting and processing massive amounts of data, which albeit beneficial, could pose the public to huge privacy, computer security, and ethical risks. Designing secure, privacy-respecting, and equitable smart city technologies requires an in-depth understanding of people's attitudes, privacy concerns, and preferences toward data collections of smart city technologies. To obtain rich qualitative data, we conducted a set of 21 semi-structured interviews with a diverse sample of participants from most underserved neighborhoods of Seattle, a large West Coast city of the United States. We complemented the interview study with a large-scale survey with 348 Prolific participants from the US. Through the interview and survey studies, we qualitatively and quantitatively measured the significance of various data collection factors, along with participants' demographic information, on their privacy preferences, concerns, and expectations toward IoT data collection scenarios in cities. Our qualitative and quantitative analyses of participants' responses surfaced several key issues that city stakeholders should consider to design safe, protective, and equitable technologies and policies.

ACKNOWLEDGMENTS

This work was supported in part by the U.S. National Science Foundation under awards CNS-1565252 and CNS-2114230, the University of Washington Tech Policy Lab (which receives support from the William and Flora Hewlett Foundation, the John D. and Catherine T. MacArthur Foundation, Microsoft, and the Pierre and Pamela Omidyar Fund at the Silicon Valley Community Foundation), and gifts from Google and Woven Planet. Wei Dai's contribution to this work was conducted during a previous employment at Microsoft.

REFERENCES

[1] Desiree Abrokwa, Shruti Das, Omer Akgul, and Michelle L Mazurek. 2021. Comparing Security and Privacy Attitudes Among US Users of Different Smartphone and Smart-Speaker Platforms. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. 139–158.

[2] Fola Akinnibi. [n. d.]. NYC, Chicago Waste Millions on Gunshot Detection Technology, Report Says. https://www.bloomberg.com/news/articles/2022-07-14/new-york-chicago-waste-millions-on-gunshot-detection-tech-privacy-group-says. Accessed: 2022-12-13.

[3] Titi Akinsanmi and Aishat Salami. 2021. Evaluating the trade-off between privacy, public health safety, and digital security in a pandemic. *Data & Policy* 3 (2021).

[4] Vito Albino, Umberto Berardi, and Rosa Maria Dangelico. 2015. Smart cities: Definitions, dimensions, performance, and initiatives. *Journal of urban technology* 22, 1 (2015), 3–21.

[5] Hamza Alshenqeeti. 2014. Interviewing as a data collection method: A critical review. *English Linguistics Research* 3, 1 (2014), 39.

[6] American Civil Liberties Union (ACLU). [n. d.]. Community Control Over Police Surveillance (CCOPS). https://www.aclu.org/issues/privacy-technology/surveillance-technologies/community-control-over-police-surveillance. Accessed: 2021-05-14.

[7] Leonidas Anthopoulos and Panos Fitsilis. [n. d.]. Considering Future Internet on the Basis of Smart Urban Cities. ([n. d.]).

[8] Katrin Arning, Sylvia Kowalewski, and Martina Ziefle. 2012. Modelling user acceptance of wireless medical technologies. In *International Conference on Wireless Mobile Communication and Healthcare*. Springer, 146–153.

[9] Katrin Arning, Martina Ziefle, and Heike Muehlhans. 2013. Join the ride! user requirements and interface design guidelines for a commuter carpooling platform. In *International Conference of Design, User Experience, and Usability*. Springer, 10–19.

[10] Brooke Auxier, Lee Rainie, Monica Anderson, Andrew Perrin, Madhu Kumar, and Erica Turner. 2019. Americans and Privacy: Concerned, Confused and Feeling Lack of Control Over Their Personal Information. https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/.

[11] Kristen Backor, Saar Golde, and Norman Nie. 2007. Estimating Survey Fatigue in Time Use Study. In *International Association for Time Use Research Conference. Washington, DC*. Citeseer. Accessed: 2020-06-03.

[12] Kirstie Ball. 2019. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. *Surveillance & Society* 17, 1/2 (2019), 252–256.

[13] BBC Future. [n. d.]. Safe cities: Using smart tech for public security. https://www.bbc.com/future/bespoke/specials/connected-world/government.html. Accessed: 2021-04-28.

[14] Daniel Belanche-Gracia, Luis V Casaló-Ariño, and Alfredo Pérez-Rueda. 2015. Determinants of multi-service smartcard success for smart cities development: A study based on citizens' privacy and security perceptions. *Government information quarterly* 32, 2 (2015), 154–163.

[15] Michael Birnhack and Lotem Perry-Hazan. 2020. School surveillance in context: High school students' perspectives on CCTV, privacy, and security. *Youth & Society* 52, 7 (2020), 1312–1330.

[16] Albert Blankenship. 1942. Psychological difficulties in measuring consumer preference. *Journal of Marketing* 6, 4_part_2 (1942), 66–75.

[17] Amanda Brandeis. 2021. As smart city technology advances, community groups warn of privacy concerns. https://www.thedenverchannel.com/news/national/as-smart-city-technology-advances-community-groups-warn-of-privacy-concern.

[18] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[19] Kenneth P Burnham and David R Anderson. 2004. Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods & research* 33, 2 (2004), 261–304.

[20] Sara Cannizzaro, Rob Procter, Sinong Ma, and Carsten Maple. 2019. Adoption and acceptability of smart devices for the home. (2019).

[21] Sara Cannizzaro, Rob Procter, Sinong Ma, and Carsten Maple. 2020. Trust in the smart home: Findings from a nationally representative survey in the UK. *Plos one* 15, 5 (2020), e0231615.

[22] Xavier Caron, Rachelle Bosua, Sean B Maynard, and Atif Ahmad. 2016. The Internet of Things (IoT) and its impact on individual privacy: An Australian perspective. *Computer law & security review* 32, 1 (2016), 4–15.

[23] Randy Carroll, Rick Cnossen, Mark Schnell, and David Simons. 2007. Continua: An interoperable personal healthcare ecosystem. *IEEE Pervasive Computing* 6, 4 (2007), 90–94.

[24] Marianna Cavada, DV Hunt, and CD Rogers. 2014. Smart cities: Contradicting definitions and unclear measures. In *World Sustainability Forum*. MDPI AG, 1–12.

[25] Jacqueline G Cavazos, P Jonathon Phillips, Carlos D Castillo, and Alice J O'Toole. 2020. Accuracy comparison across face recognition algorithms: Where are we on measuring race bias? *IEEE Transactions on Biometrics, Behavior, and Identity Science* 3, 1 (2020), 101–111.

[26] Leyland Cecco. 2019. 'Surveillance capitalism': critic urges Toronto to abandon smart city project. https://www.theguardian.com/cities/2019/jun/06/toronto-smart-city-google-project-privacy-concerns.

[27] Sheshadri Chatterjee, Arpan Kumar Kar, and MP Gupta. 2018. Alignment of IT authority and citizens of proposed smart cities in India: System security and privacy perspective. *Global Journal of Flexible Systems Management* 19, 1 (2018), 95–107.

[28] Brian Chen. [n. d.]. Security Cameras Make Us Feel Safe, but Are They Worth the Invasion? https://www.nytimes.com/2022/11/02/technology/personaltech/security-cameras-surveillance-privacy.html. Accessed: 2022-12-13.

[29] Hafedh Chourabi, Taewoo Nam, Shawn Walker, J Ramon Gil-Garcia, Sehl Mellouli, Karine Nahon, Theresa A Pardo, and Hans Jochen Scholl. 2012. Understanding smart cities: An integrative framework. In *2012 45th Hawaii international conference on system sciences*. IEEE, 2289–2297.

[30] Richard Chow. 2017. The last mile for IoT privacy. *IEEE Security & Privacy* 15, 6 (2017), 73–76.

[31] Liezel Cilliers and Stephen Flowerday. 2014. Information security in a public safety, participatory crowdsourcing smart city project. In *World Congress on Internet Security (WorldCIS-2014)*. IEEE, 36–41.

[32] Liezel Cilliers and Stephen Flowerday. 2015. The Relationship Between Privacy, Information Security and the Trustworthiness of a Crowdsourcing System in a Smart City. In *HAISA*. 243–255.

[33] James Clayton. [n. d.]. Inside the controversial US gunshot-detection firm. https://www.bbc.com/news/technology-59072745. Accessed: 2022-12-13.

[34] Eric Corbett and Christopher Le Dantec. 2021. Designing Civic Technology with Trust. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–17.

[35] Eric Corbett and Christopher A Le Dantec. 2018. The problem of community engagement: Disentangling the practices of municipal government. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.

[36] Cybersecurity and Infrastructure Security Agency. [n. d.]. Trust In Smart City Systems Report. https://www.cisa.gov/publication/trust-smart-city-systems-report. Accessed: 2022-12-13.

[37] Corey Dickens, Paul Boynton, and Sokwoo Rhee. 2019. Principles for designed-in security and privacy for smart cities. In *Proceedings of the Fourth Workshop on International Science of Smart City Operations and Platforms Engineering*. 25–29.

[38] Jessa Dickinson, Mark Díaz, Christopher A Le Dantec, and Sheena Erete. 2019. " The cavalry ain't coming in to save us" Supporting Capacities and Relationships through Civic Tech. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–21.

[39] Hector Dominguez, Judith Mowry, Elisabeth Perez, Christine Kendrick, and Kevin Martin. 2019. Privacy and information protection for a new generation of city services. In *Proceedings of the 2nd ACM/EIGSCC Symposium on Smart Cities and Communities*. 1–6.

[40] Mitchell L Doucette, Christa Green, Jennifer Necci Dineen, David Shapiro, and Kerri M Raissian. 2021. Impact of ShotSpotter Technology on Firearm Homicides and Arrests Among Large Metropolitan Counties: A Longitudinal Analysis, 1999–2016. *Journal of urban health* 98, 5 (2021), 609–621.

[41] Pardis Emami Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Faith Cranor, and Norman Sadeh. 2017. Privacy expectations and preferences in an IoT world. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. 399–412.

[42] Pardis Emami Naeini, Martin Degeling, Lujo Bauer, Richard Chow, Lorrie Faith Cranor, Mohammad Reza Haghighat, and Heather Patterson. 2018. The influence of friends and experts on privacy decision making in IoT scenarios. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–26.

[43] Pardis Emami-Naeini, Henry Dixon, Yuvraj Agarwal, and Lorrie Faith Cranor. 2019. Exploring how privacy and security factor into IoT device purchase behavior. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[44] Pardis Emami-Naeini, Tiona Francisco, Tadayoshi Kohno, and Franziska Roesner. 2021. Understanding Privacy Attitudes and Concerns Towards Remote Communications During the COVID-19 Pandemic. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. 695–714.

[45] Fariborz Farahmand and Eugene H Spafford. 2013. Understanding insiders: An analysis of risk-taking behavior. *Information systems frontiers* 15, 1 (2013), 5–15.

[46] Alessandra Fasano. 2014. Smart City: Fact and Fiction. *Available at SSRN 2410739* (2014).

[47] Teri Figueroa. 2020. Mayor orders San Diego's Smart Streetlights turned off until surveillance ordinance in place. https://www.sandiegouniontribune.com/news/public-safety/story/2020-09-09/mayor-orders-san-diegos-smart-streetlights-turned-off-until-surveillance-ordinance-in-place.

[48] Teri Figueroa and Lori Weisberg. 2020. San Diego to Consider New Smart Streetlight Privacy Policy. https://www.govtech.com/computing/san-diego-to-

consider-new-smart-streetlight-privacy-policy.html.
[49] Janet Finch. 1987. The vignette technique in survey research. *Sociology* 21, 1 (1987), 105–114.
[50] Jennifer Gabrys. 2014. Programming environments: Environmentality and citizen sensing in the smart city. *Environment and Planning D: Society and Space* 32, 1 (2014), 30–48.
[51] Vaibhav Garg and Jean Camp. 2012. End user perception of online risk under uncertainty. In *2012 45th Hawaii International Conference on System Sciences*. IEEE, 3278–3287.
[52] Michele Gilman and Rebecca Green. 2018. The surveillance gap: The harms of extreme privacy and data marginalization. *NYU Rev. L. & Soc. Change* 42 (2018), 253.
[53] GlobalData Thematic Research. 2020. History of smart cities: Timeline. https://www.verdict.co.uk/smart-cities-timeline.
[54] Daniel Gooch, Matthew Barker, Lorraine Hudson, Ryan Kelly, Gerd Kortuem, Janet Van Der Linden, Marian Petre, Rebecca Brown, Anna Klis-Davies, Hannah Forbes, et al. 2018. Amplifying quiet voices: Challenges and opportunities for participatory design at an urban scale. *ACM Transactions on Computer-Human Interaction (TOCHI)* 25, 1 (2018), 1–34.
[55] Daniel Gooch, Ryan M Kelly, Alexandra Stiver, Janet van der Linden, Marian Petre, Mike Richards, Anna Klis-Davies, Jessica MacKinnon, Robbie Macpherson, and Clare Walton. 2020. The benefits and challenges of using crowdfunding to facilitate community-led projects in the context of digital civics. *International Journal of Human-Computer Studies* 134 (2020), 33–43.
[56] Tirso Javier Hernández Gracia, Alejandra Corichi García, et al. 2018. Sustainable smart cities. creating spaces for technological, social and business development. *Boletín Científico de Las Ciencias Económico Administrativas Del ICEA* 6, 12 (2018).
[57] Christina Gross, Markus Siepermann, and Richard Lackes. 2020. The acceptance of smart home technology. In *International Conference on Business Informatics Research*. Springer, 3–18.
[58] Matthew Guariglia. [n. d.]. It's Time for Police to Stop Using ShotSpotter. https://www.eff.org/deeplinks/2021/07/its-time-police-stop-using-shotspotter. Accessed: 2022-12-13.
[59] Gary Gumpert and Susan J Drucker. 2001. Public boundaries: Privacy and surveillance in a technological world. *Communication Quarterly* 49, 2 (2001), 115–129.
[60] Colin Harrison, Barbara Eckman, Rick Hamilton, Perry Hartswick, Jayant Kalagnanam, Jurij Paraszczak, and Peter Williams. 2010. Foundations for smarter cities. *IBM Journal of research and development* 54, 4 (2010), 1–16.
[61] Line Kristin Haug. 2019. *Is Smart city sustainable? A case study of Bodø municipality*. Master's thesis. Nord universitet.
[62] Andrew Hawkins. 2020. The ACLU is suing Los Angeles over its controversial scooter tracking system. https://www.theverge.com/2020/6/8/21284490/aclu-ladot-mds-lawsuit-scooter-tracking-uber.
[63] Andrew Hawkins. 2020. Alphabet's Sidewalk Labs shuts down Toronto smart city project. https://www.theverge.com/2020/5/7/21250594/alphabet-sidewalk-labs-toronto-quayside-shutting-down.
[64] Dirk Helbing. 2021. Summary: What's Wrong with AI? In *Next Civilization*. Springer, 285–313.
[65] Dirk Helbing, Farzam Fanitabasi, Fosca Giannotti, Regula Hänggli, Carina I Hausladen, Jeroen van den Hoven, Sachit Mahajan, Dino Pedreschi, and Evangelos Pournaras. 2021. Ethics of smart cities: Towards value-sensitive design and co-evolving city life. *Sustainability* 13, 20 (2021), 11162.
[66] Drew Hemment and Anthony Townsend. 2013. Smart citizens. (2013).
[67] Sarah Holder. 2018. The Shadowy Side of LED Streetlights. https://www.bloomberg.com/news/articles/2018-03-08/led-streetlights-raise-new-opportunities-for-surveillance.
[68] Lorraine Hudson, Annika Wolff, Daniel Gooch, Janet Van Der Linden, Gerd Kortuem, Marian Petre, Rianne ten Veen, and Sinead O'Connor-Gotra. 2019. Supporting urban change: Using a MOOC to facilitate attitudinal learning and participation in smart cities. *Computers & Education* 129 (2019), 37–47.
[69] Trevor Hunter, Peter Worthy, Stephen Viller, and Ben Matthews. 2018. Designing for citizen-centric smart cities: supporting people's needs of a community garden. In *Proceedings of the 30th Australian Conference on Computer-Human Interaction*. 162–166.
[70] IESE Business School. [n. d.]. These Are The 10 Smartest Cities In The World For 2020. https://www.forbes.com/sites/iese/2020/07/08/these-are-the-10-smartest-cities-in-the-world-for-2020/?sh=35c2267212af. Accessed: 2021-04-28.
[71] Jann Ingmire. [n. d.]. Young people of color mistrust police and legal system, report finds. https://news.uchicago.edu/story/young-people-color-mistrust-police-and-legal-system-report-finds. Accessed: 2022-12-13.
[72] Intelligent Transport. 2021. Seattle unveils ambitious green transportation plan. https://www.intelligenttransport.com/transport-news/120010/seattle-climate-plan/.
[73] Athina Ioannou and Iis Tussyadiah. 2021. Privacy and surveillance attitudes during health crises: Acceptance of surveillance and privacy protection behaviours. *Technology in Society* 67 (2021), 101774.
[74] Elvira Ismagilova, Laurie Hughes, Nripendra P Rana, and Yogesh K Dwivedi. 2020. Security, privacy and risks within smart cities: Literature review and development of a smart city interaction framework. *Information Systems Frontiers* (2020), 1–22.
[75] Ryan Johnston. 2020. Digital privacy concerns will follow Sidewalk Labs to next venture, says former consultant. https://statescoop.com/ann-cavoukian-waterfront-toronto-digital-privacy-concerns-sidewalk-labs/.
[76] Joseph B Kadane and Nicole A Lazar. 2004. Methods and Criteria for Model Selection. *J. Amer. Statist. Assoc.* 99, 465 (2004), 279–290. https://doi.org/10.1198/016214504000000269 arXiv:https://doi.org/10.1198/016214504000000269
[77] Rob Kitchin. 2016. The ethics of smart cities and urban science. *Philosophical transactions of the royal society A: Mathematical, physical and engineering sciences* 374, 2083 (2016), 20160115.
[78] Francisco R Klauser and Anders Albrechtslund. 2014. From self-tracking to smart urban infrastructures: Towards an interdisciplinary research agenda on Big Data. *Surveillance & Society* 12, 2 (2014), 273–286.
[79] SN Kondepudi, V Ramanarayanan, A Jain, GN Singh, NK Nitin Agarwal, R Kumar, R Singh, P Bergmark, T Hashitani, P Gemma, et al. 2014. Smart sustainable cities analysis of definitions. *The ITU-T focus group for smart sustainable cities* (2014).
[80] Rodger James Lea. 2017. Smart cities: An overview of the technology trends driving smart cities. (2017).
[81] Hosub Lee and Alfred Kobsa. 2016. Understanding user privacy in Internet of Things environments. In *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)*. IEEE, 407–412.
[82] Hosub Lee and Alfred Kobsa. 2017. Privacy preference modeling and prediction in a simulated campuswide IoT environment. In *2017 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 276–285.
[83] LinkNYC. [n. d.]. Free super fast Wi-Fi. And that's just the beginning. https://www.link.nyc/. Accessed: 2021-04-25.
[84] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhimedi, Shikun Aerin Zhang, Norman Sadeh, Yuvraj Agarwal, and Alessandro Acquisti. 2016. Follow my recommendations: A personalized privacy assistant for mobile app permissions. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*. 27–41.
[85] Kathleen M Macqueen, Eleanor McLellan-Lemal, Kelly Bartholow, and Bobby Milstein. 2008. Team-based codebook development: Structure, process, and agreement. *Handbook for team-based qualitative research* (2008), 119–135.
[86] Teena Maddox. 2016. Gunshot detection technology as part of smart city design. https://www.techrepublic.com/article/gunshot-detection-technology-as-part-of-smart-city-design/.
[87] Ryan Mark and Gregory Anya. 2019. Ethics of using smart city AI and big data: The case of four large European cities. *The ORBIT Journal* 2, 2 (2019), 1–36.
[88] Kirsten Martin and Helen Nissenbaum. 2016. Measuring privacy: An empirical test using context to expose confounding variables. *Colum. Sci. & Tech. L. Rev.* 18 (2016), 176.
[89] Kirsten E Martin. 2012. Diminished or just different? A factorial vignette study of privacy as a social contract. *Journal of Business Ethics* 111, 4 (2012), 519–539.
[90] Antoni Martínez-Ballesté, Pablo A Pérez-Martínez, and Agusti Solanas. 2013. The pursuit of citizens' privacy: a privacy-aware smart city is possible. *IEEE Communications Magazine* 51, 6 (2013), 136–141.
[91] Gary T Marx. 1998. Ethics for the new surveillance. *The Information Society* 14, 3 (1998), 171–185.
[92] Jesse Marx. 2019. Public, Council Were in the Dark on Police Access to 'Smart' Streetlights. https://www.voiceofsandiego.org/topics/public-safety/public-council-were-in-the-dark-on-police-access-to-smart-streetlights/.
[93] Lorraine Green Mazerolle. 1998. *Using gunshot detection technology in high-crime areas*. US Department of Justice, Office of Justice Programs, National Institute of ....
[94] H Patricia McKenna. 2016. Is it all about awareness? People, smart cities 3.0, and evolving spaces for IT. In *Proceedings of the 2016 ACM SIGMIS Conference on Computers and People Research*. 47–56.
[95] Lucas Melgaço and Rosamunde van Brakel. 2021. Smart cities as surveillance theatre. *Surveillance & Society* 19, 2 (2021), 244–249.
[96] Martha Mendoza and Garance Burke. [n. d.]. How AI-powered tech landed man in jail with scant evidence. https://apnews.com/article/artificial-intelligence-algorithm-technology-police-crime-7e3345485aa668c97606d4b54f9b6220. Accessed: 2022-09-13.
[97] Merriam-Webster. [n. d.]. Privacy. https://www.merriam-webster.com/dictionary/privacy. Accessed: 2021-04-28.
[98] Saraju P Mohanty, Uma Choppali, and Elias Kougianos. 2016. Everything you wanted to know about smart cities: The internet of things is the backbone. *IEEE Consumer Electronics Magazine* 5, 3 (2016), 60–70.
[99] David Murakami Wood. [n. d.]. Smart City, Surveillance City. https://www.scl.org/articles/3405-smart-city-surveillance-city. Accessed: 2022-12-13.
[100] Taewoo Nam and Theresa A Pardo. 2011. Conceptualizing smart city with dimensions of technology, people, and institutions. In *Proceedings of the 12th annual international digital government research conference: digital government*

innovation in challenging times. 282–291.
[101] Taewoo Nam and Theresa A Pardo. 2011. Smart city as urban innovation: Focusing on management, policy, and context. In *Proceedings of the 5th international conference on theory and practice of electronic governance.* 185–194.
[102] Brian Nordli. 2020. How the Array of Things Project is Making Chicago a Smart City. https://builtin.com/founders-entrepreneurship/smart-city-chicago-array-of-things-project.
[103] Kieron O'Hara, MH C Nguyen, and Peter Haynes. 2014. *Digital Enlightenment Yearbook 2014: Social Networks and Social Machines, Surveillance and Empowerment.* IOS Press.
[104] Paul A Pavlou. 2011. State of the information privacy literature: Where are we now and where should we go? *MIS quarterly* (2011), 977–988.
[105] Gloria Pazmino. 2019. There are Surveillance Questions About LinkNYC Kiosks — and They're Losing Money. https://www.ny1.com/nyc/all-boroughs/politics/2019/05/09/linknyc-kiosks-surveillance-questions-about-cameras-recording-kiosks-also-losing-money.
[106] Eyal Peer, David M Rothschild, Zak Evernden, Andrew Gordon, and Ekaterina Damer. 2021. MTurk, Prolific or panels? Choosing the right audience for online research. *Choosing the right audience for online research (January 10, 2021)* (2021).
[107] Stephen R Porter, Michael E Whitcomb, and William H Weitzer. 2004. Multiple surveys of students and survey fatigue. *New directions for institutional research* 2004, 121 (2004), 63–73.
[108] Blaine A Price, Karim Adam, and Bashar Nuseibeh. 2005. Keeping ubiquitous computing to yourself: A practical model for user control of privacy. *International Journal of Human-Computer Studies* 63, 1-2 (2005), 228–253.
[109] Prolific. 2020. Why Prolific? https://www.prolific.co/prolific-vs-mturk/. Accessed: 2021-07-06.
[110] Catharine H Rankin, Thomas Abrams, Robert J Barry, Seema Bhatnagar, David F Clayton, John Colombo, Gianluca Coppola, Mark A Geyer, David L Glanzman, Stephen Marsland, et al. 2009. Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. *Neurobiology of learning and memory* 92, 2 (2009), 135–138.
[111] Retail Sensing. 2020. Smart City Solutions. https://www.retailsensing.com/smart-city-iot.html.
[112] Norman Sadeh, Bin Liu, Anupam Das, Martin Degeling, and Florian Schaub. 2021. Personalized privacy assistant. US Patent 10,956,586.
[113] Johnny Saldaña. 2021. *The coding manual for qualitative researchers.* sage.
[114] Laura Santhanam. [n. d.]. Two-thirds of black Americans don't trust the police to treat them equally. Most white Americans do. https://www.pbs.org/newshour/politics/two-thirds-of-black-americans-dont-trust-the-police-to-treat-them-equally-most-white-americans-do. Accessed: 2022-12-13.
[115] Taylor Shelton, Matthew Zook, and Alan Wiig. 2015. The 'actually existing smart city'. *Cambridge journal of regions, economy and society* 8, 1 (2015), 13–25.
[116] Lucy Simko, Ryan Calo, Franziska Roesner, and Tadayoshi Kohno. 2020. COVID-19 contact tracing and privacy: studying opinion and preferences. *arXiv preprint arXiv:2005.06056* (2020).
[117] Jay Stanley. [n. d.]. Four Problems with the ShotSpotter Gunshot Detection System. https://www.aclu.org/news/privacy-technology/four-problems-with-the-shotspotter-gunshot-detection-system. Accessed: 2022-12-13.
[118] Katy Stegall. 2020. Streetlight Spy Cameras Have Led to a Massive Privacy Backlash in San Diego. https://www.vice.com/en/article/z3vn83/streetlight-spy-cameras-have-led-to-a-massive-privacy-backlash-in-san-diego.
[119] Linnet Taylor and Christine Richter. 2015. Big data and urban governance. In *Geographies of urban governance.* Springer, 175–191.
[120] Cristian Toma, Andrei Alexandru, Marius Popa, and Alin Zamfiroiu. 2019. IoT solution for smart cities' pollution monitoring and the security challenges. *Sensors* 19, 15 (2019), 3401.
[121] Nicol Turner Lee and Caitlin Chin. [n. d.]. Police surveillance and facial recognition: Why data privacy is imperative for communities of color. https://www.brookings.edu/research/police-surveillance-and-facial-recognition-why-data-privacy-is-an-imperative-for-communities-of-color/. Accessed: 2022-12-13.
[122] United Nations, Department of Economic and Social Affairs. [n. d.]. 68% of the world population projected to live in urban areas by 2050, says UN. https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html. Accessed: 2021-04-24.
[123] Julia van Heek, Katrin Aming, and Martina Ziefle. 2016. "How fear of crime affects needs for privacy & safety": Acceptance of surveillance technologies in smart cities. In *2016 5th International Conference on Smart Cities and Green ICT Systems (SMARTGREENS).* IEEE, 1–12.
[124] Liesbet Van Zoonen. 2016. Privacy concerns in smart cities. *Government Information Quarterly* 33, 3 (2016), 472–480.
[125] Sanjana Kadaba Viswanath, Sai Ram Gubba, Balasundram Arunn, Chandra Sekar Veerappan, and Chau Yuen. 2014. On the design of a cost-effective and light-weight people counting sensor. In *International Internet of Things Summit.* Springer, 176–182.
[126] Jane Wakefield. [n. d.]. Google-linked smart city plan ditched in Portland. https://www.bbc.com/news/technology-56168306. Accessed: 2022-12-13.
[127] Junjue Wang, Brandon Amos, Anupam Das, Padmanabhan Pillai, Norman Sadeh, and Mahadev Satyanarayanan. 2017. A scalable and privacy-aware IoT service for live video analytics. In *Proceedings of the 8th ACM on Multimedia Systems Conference.* 38–49.
[128] Sheng-Ming Wang and Chieh-Ju Huang. 2014. User experience analysis on urban interaction and information service in smart city nodes. In *Proceedings of the Second International Symposium of Chinese CHI.* 103–109.
[129] Yang Wang, Yun Huang, and Claudia Louis. 2013. Respecting user privacy in mobile crowdsourcing. *Science* 2, 2 (2013), pp–50.
[130] Mika Westerlund, Diane A Isabelle, and Seppo Leminen. 2021. The acceptance of digital surveillance in an age of big data. (2021).
[131] Reginald Whitaker. 1998. *The end of privacy: How total surveillance is becoming a reality.* New Press.
[132] Cedric Deslandes Whitney, Teresa Naval, Elizabeth Quepons, Simrandeep Singh, Steven R Rick, and Lilly Irani. 2021. HCI Tactics for Politics from Below: Meeting the Challenges of Smart Cities. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems.* 1–15.
[133] Wikipedia. [n. d.]. Smart city. https://en.wikipedia.org/wiki/Smart_city. Accessed: 2021-04-21.
[134] David Murakami Wood and Valerie Steeves. 2021. Smart Surveillance. *Surveillance & Society* 19, 2 (2021), 150–153.
[135] Nesrine Ben Yahia, Wissem Eljaoued, Narjès Bellamine Ben Saoud, and Ricardo Colomo-Palacios. 2019. Towards sustainable collaborative networks for smart cities co-governance. *International Journal of Information Management* (2019), 102037.
[136] Wenxuan Yu and Chengwei Xu. 2018. Developing smart cities in China: An empirical analysis. *International Journal of Public Administration in the Digital Age (IJPADA)* 5, 3 (2018), 76–91.
[137] Yuejin Zhang, Fangyao Liu, Zhiqiang Gu, Zhengxin Chen, Yong Shi, and Aihua Li. 2019. Research on smart city evaluation based on hierarchy of needs. *Procedia Computer Science* 162 (2019), 467–474.
[138] Jan Henrik Ziegeldorf, Oscar Garcia Morchon, and Klaus Wehrle. 2014. Privacy in the Internet of Things: threats and challenges. *Security and Communication Networks* 7, 12 (2014), 2728–2742.

A INTERVIEW PROCEDURE

A.1 Screening Survey

A.1.1 Informed Consent. We are a team of researchers at the University of Washington in Seattle. In this survey, we will ask you some demographic questions. If you are eligible to participate in our main interview study, we will email you in the next few days to schedule the remote interview session. The Human Subjects Division at the University of Washington reviewed our study, and determined that it was exempt from federal human subjects regulation. We do not expect that this survey will put you at any risk for harm.

In order to participate, you must be at least 18 years old and able to complete the survey in English. We expect this survey will take about 5 minutes to complete. If you have any questions about this survey, you may email us at pardis@cs.washington.edu.

- I am 18 years or older.
  ○ Yes ○ No
- I have read and understand the information above.
  ○ Yes ○ No
- I want to participate in this research and continue with the task.
  ○ Yes ○ No

A.1.2 Screening Questions.

- What is your age? [Open-ended]
- What is your gender? [Open-ended]
- What is the highest degree you have earned?
  ○ No schooling completed ○ Nursery school ○ Grades 1 through 11 ○ 12th grade—no diploma ○ Regular high school diploma ○ GED or alternative credential ○ Some college credit, but less than 1 year of college ○ 1 or more years of college credit, no degree ○ Associate's degree (for example: AA, AS) ○ Bachelor's degree (for example: BA, BS) ○ Master's degree (for example: MA, MS, MEng, MEd, MSW, MBA) ○ Professional degree beyond bachelor's degree (for example: MD, DDS, DVM, LLB, JD) ○ Doctorate degree (for example: PhD, EdD)
- What is your current employment status?
  ○ Full-time employment ○ Part-time employment ○ Unemployed ○ Self-employed ○ Home-maker ○ Student ○ Retired
- What was your most recent employment status before the COVID-19 pandemic?
  ○ Full-time employment ○ Part-time employment ○ Unemployed ○ Self-employed ○ Home-maker ○ Student ○ Retired
- *(If Unemployed is not selected)* Before the COVID-19 pandemic, what was your primary mode of transportation for a typical trip to work? By primary, we mean the modes of transportation you used for the longest portion of the trip (Check as many as apply).
  ○ Bicycle ○ Car or truck (solo commute) ○ Car pool or van pool ○ Motorcycle ○ Primarily worked from home ○ Public transportation (e.g., bus, rail) ○ Taxi or rideshare (e.g., Uber, Lyft) ○ Walking ○ Other (please specify [Open-ended])
- In which neighborhood of Seattle do you live in? [Options were the neighborhoods of the city we recruited from.]
- Do you have a background in technology?
  ○ Yes ○ No
- *(If Yes is selected)* Please specify what your technical background is. [Open-ended]
- Please specify your entire yearly household income (in 2020) before taxes.
  ○ Less than $10,000 ○ $10,000 to $19,999 ○ $20,000 to $29,999 ○ $30,000 to $39,999 ○ $40,000 to $49,999 ○ $50,000 to $59,999 ○ $60,000 to $69,999 ○ $70,000 to $79,999 ○ $80,000 to $89,999 ○ $90,000 to $99,999 ○ $100,000 to $149,999 ○ $150,000 or more
- Including yourself, how many adults 18 years of age and above live in your current home?
  ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ More than 5
- How many children under the age of 18 live in your current home?
  ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ More than 5
- Who do you share your home with? (check as many as apply)
  ○ No one ○ Roommate(s) ○ Spouse(s)/Domestic partner(s) ○ Children ○ Parent(s) ○ Other (please specify [Open-ended])
- Before the COVID-19 pandemic, how many hours a week did you spend outside of your home?
  ○ Less than 1 ○ 1 to 5 hours ○ 6 to 10 hours ○ 11 to 15 hours ○ 16 to 20 hours ○ 21 to 25 hours ○ 26 to 30 hours ○ 31 to 35 hours ○ 36 to 40 hours ○ Over 40 hours
- How do you describe your ethnicity? (Check as many as apply)
  ○ American Indian or Alaskan Native ○ Asian ○ Black or African American ○ Hispanic or Latino or Spanish Origin of any race ○ Native Hawaiian or Other Pacific Islander ○ White ○ Other (please specify [Open-ended])
- How well do you speak English?
  ○ Very well ○ Well ○ Not well ○ Not at all
- For the interview, we need you to have access to a smartphone, a laptop, a tablet, or a desktop computer. What application would you prefer to use for the remote interview session? (Check as many as apply)
  ○ Regular phone call ○ FaceTime ○ Google Hangouts ○ Microsoft Teams ○ Skype ○ Zoom ○ Other (please specify [Open-ended])
- Where did you hear about our study?
  ○ Posting on Craigslist ○ Posting on a location other than Craigslist ○ Other (please specify [Open-ended])
- If you are eligible to participate in our interview study, what is the best email address that we can use to contact you to schedule the interview session? [Open-ended]

A.2 Interview Questions

A.2.1 Knowledge and Awareness about a Smart City.

- Have you ever heard of a smart city?
- *(If the answer is Yes)* In your own words, how would you define a smart city?

- *(If the answer is No)* If you were to guess, how would you define a smart city?

(We provide a definition of smart city for all participants) A smart city has different definitions. In one of the closely-related definitions, a smart city is described as an urban area that uses different types of electronic methods and sensors to collect data. Insights gained from that data are used to manage assets, resources, and services efficiently.

- Do you know of any cities which are leveraging such a smart sensing system?
- *(If the answer is Yes)* What information is being collected about the city itself or the people who live in the city?
- *(If the answer is Yes)* How did you hear, read, or learn about this information?
- What information do you think the city should know about the people who live in the city and why?
- What information do you think the city should not know about the people who live in the city and why?

A.2.2 Attitudes toward Smart City Data Collection and Use Scenarios. By changing the levels of three factors of `data_type` *(with four levels),* `data_access` *(with four levels), and* `data_retention` *(with two levels), we presented eight smart city vignettes to participants. After presenting each scenario, we asked follow up questions. Below is an example scenario that we showed to participants.*

> Imagine you are walking in the street where your home is located in Seattle, and you see a streetlamp. There are sensors on this streetlamp, which can only collect **real-time video footage of people** in the proximity of the streetlamp. **Everybody** can access this collected information. The collected information will **never get deleted**.

- Do you have any questions about the scenario that we just described?
- Do you think the described scenario is currently happening in any city?
- *(If the answer is Yes)* What cities are currently collecting this information?
- *(If the answer is No)* Why do you think this scenario is not currently happening?
- How concerned are you with the described scenario?
- *(If the participant is concerned about the described scenario)* What specifically about this scenario makes you concerned?
- *(If the participant is concerned about the described scenario)* What do you think should happen to make you less concerned with this scenario?
- What are some potential benefits and harms associated with this scenario?
- How interested would you be in being informed about such information collections and use? how often and Why?

A.2.3 Final Questions.

- Do you think other people in your neighborhood would have perceived the benefits and harms related to the smart city information collection and use scenarios the same way or a different way than you? Why?
- How different or similar do you imagine yourself thinking about the smart city information collection and use scenarios after this interview versus before the interview?

B SURVEY PROCEDURE

We presented each participant with four IoT data collection scenarios and asked the same questions after each scenario. Here we will provide one example scenario and its follow-up questions.

B.1 Informed Consent

This is a survey about the use of technologies in cities by researchers at the University of Washington.

Our institution's Human Subjects Division reviewed our study and determined that it was exempt from federal human subjects regulation. We do not expect that this survey will put you at any risk for harm.

To participate, you must be at least 18 years old and able to complete the survey in English. We expect this survey will take about 15 minutes to complete. If you have any questions about this survey, you may email us at: pardis@cs.washington.edu.

- I am 18 years or older.
  ○ Yes ○ No
- I have read and understand the information above.
  ○ Yes ○ No
- I want to participate in this research and continue with the task.
  ○ Yes ○ No

B.2 Introduction

In this survey, we will show you four data collection scenarios. After presenting each scenario, we will ask you some follow-up questions. We will end the survey by asking general demographic questions.

B.3 Scenario and Follow-Up Questions

> *Imagine you are walking in the street where your home is located and you see a streetlamp. There are sensors on this streetlamp, which can only collect real-time video footage of people in the proximity of the streetlamp. Everybody can access this collected information. The collected information will never get deleted.*

- How concerned are you about this described data collection scenario?
  ○ Extremely concerned ○ Moderately concerned ○ Somewhat concerned ○ Slightly concerned ○ Not at all concerned
- *(If not at all concerned is not selected)* What about this described data collection scenario makes you concerned? [Open-ended]
- *(If not at all concerned is selected)* What about this described data collection scenario makes you not at all concerned? [Open-ended]
- How beneficial do you think this described data collection scenario would be to you?
  ○ Extremely beneficial ○ Moderately beneficial ○ Somewhat beneficial ○ Slightly beneficial ○ Not at all beneficial

- *(If not at all beneficial is not selected)* How do you think this described data collection scenario would be beneficial to you? [Open-ended]
- *(If not at all beneficial is selected)* How do you think this described data collection scenario would be not at all beneficial to you?
- How beneficial do you think this described data collection scenario would be to society?
  ∘ Extremely beneficial ∘ Moderately beneficial ∘ Somewhat beneficial ∘ Slightly beneficial ∘ Not at all beneficial
- *(If not at all beneficial is not selected)* How do you think this described data collection scenario would be beneficial to society? [Open-ended]
- *(If not at all beneficial is selected)* How do you think this described data collection scenario would be not at all beneficial to society?
- How harmful do you think this described data collection scenario would be to you?
  ∘ Extremely harmful ∘ Moderately harmful ∘ Somewhat harmful ∘ Slightly harmful ∘ Not at all harmful
- *(If not at all harmful is not selected)* How do you think this described data collection scenario would be harmful to you? [Open-ended]
- *(If not at all harmful is selected)* How do you think this described data collection scenario would be not at all harmful to you?
- How harmful do you think this described data collection scenario would be to society?
  ∘ Extremely harmful ∘ Moderately harmful ∘ Somewhat harmful ∘ Slightly harmful ∘ Not at all harmful
- *(If not at all harmful is not selected)* How do you think this described data collection scenario would be harmful to society? [Open-ended]
- *(If not at all harmful is selected)* How do you think this described data collection scenario would be not at all harmful to society?
- How frequently would you like to be notified about this described data collection scenario?
  ∘ Every time this data is being collected about you ∘ Every once in a while, when this data is being collected about you ∘ Only the first time this data is being collected about you ∘ Never
- *(If never is not selected)* What about this data collection scenario would you like to be notified about the most and why? [Open-ended]
- *(If never is selected)* Why would you prefer not to be notified about this data collection scenario? [Open-ended]

B.4 Demographic Questions

We used the same demographic questions as in the interview screening survey (see Appendix A.1.2).

C INTERVIEW AND SURVEY CODEBOOK

The codebook is available at:
https://github.com/pemamina/CHI23_SmartCity_Codebook/blob/main/CHI23_SmartCity_Codebook.pdf.

D PROOF OF (3)

Given (1), the probability of being concerned given a categorical factor f for a typical participant can be written as

$$\Pr(\text{being concerned} \mid f) = \Pr(Y \geq 2 \mid f) = 1 - \Pr(Y = 1 \mid f) \overset{(a)}{=} 1 - \sigma\left(\alpha_{1|2} - \beta_{\texttt{income}} \cdot \texttt{income} - \beta_f\right), \tag{4}$$

where in (a), we freeze the typical participant random effect at its mean, i.e., zero. Letting $\alpha'_{1|2} = \alpha_{1|2} - \beta_{\texttt{income}} \cdot \texttt{income}$, combining (4) with (2) implies that the odds ratio of increased concern level is given by

$$\begin{aligned} \mathrm{OR}^{f}_{\text{concern}} &= \left(\frac{1-\sigma\left(\alpha'_{1|2}-\beta_f\right)}{\sigma\left(\alpha'_{1|2}-\beta_f\right)}\right) \Big/ \left(\frac{1-\sigma\left(\alpha'_{1|2}-\beta_{f_{\text{baseline}}}\right)}{\sigma\left(\alpha'_{1|2}-\beta_{f_{\text{baseline}}}\right)}\right) \\ &\overset{(b)}{=} \frac{\exp\left(\beta_f - \alpha'_{1|2}\right)}{\exp\left(\beta_{f_{\text{baseline}}} - \alpha'_{1|2}\right)} \\ &\overset{(c)}{=} \exp\left(\beta_f\right), \end{aligned} \tag{5}$$

where (b) follows from the definition of the sigmoid function, and (c) holds because the CLMM coefficient corresponding to the factor baseline is zero, i.e., $\beta_{f_{\text{baseline}}} = 0$. This completes the proof. The proof for numeric factors follows the same lines as above and is omitted for brevity.

E PARTICIPANTS' DEMOGRAPHIC INFORMATION

Interviewee ID	Highest Degree	Employment	Commute Mode	Income	# Adults in Home	# Children in Home	Housemate(s)
I-P1	1+ years of college (No degree)	Full-time	Car/Truck	\$70K – \$80K	2	1	Children
I-P2	1+ years of college (No degree)	Full-time	Public/Walking	\$30K – \$40K	1	0	-
I-P3	Bachelor's degree	Unemployed	Bicycle	\$20K – \$30K	4	0	Spouse/Partner
I-P4	Bachelor's degree	Self-employed	Public/Rideshare	\$30K – \$40K	1	2	Children
I-P5	Bachelor's degree	Full-time	-	<\$10K	1	0	-
I-P6	Bachelor's degree	Student	Bicycle/Public/Walking	\$50K – \$60K	1	0	-
I-P7	Professional degree	Part-time	Car/Truck/WFH/Public/Walking	\$30K – \$40K	1	0	-
I-P8	1+ years of college (No degree)	Full-time	Car/Truck	\$50K – \$60K	2	0	Spouse/Partner
I-P9	Master's degree	Self-employed	Bicycle/Public	\$100K – \$150K	4	0	Roommates
I-P10	Bachelor's degree	Unemployed	Public	\$60K – \$70K	1	0	Other
I-P11	Bachelor's degree	Self-employed	Car/Truck/Motorcycle	\$30K – \$40K	1	0	-
I-P12	Bachelor's degree	Part-time	Car/Truck/Public/Walking	\$70K – \$80K	2	0	Spouse/Partner
I-P13	1+ years of college (No degree)	Full-time	Car/Truck	\$30K – \$40K	1	2	Children
I-P14	GED or alternative credential	Full-time	Public	\$30K – \$40K	2	0	Roommates
I-P15	Bachelor's degree	Part-time	Public	\$20K – \$30K	1	0	-
I-P16	Master's degree	Full-time	Carpool/Public/Walking	\$40K – \$50K	1	0	-
I-P17	Bachelor's degree	Retired	Car/Truck	\$10K – \$20K	2	0	Spouse/Partner
I-P18	Master's degree	Home-maker	Other	\$90K – \$100K	2	2	Spouse/Partner
I-P19	Bachelor's degree	Unemployed	Public/Rideshare/Walking	\$60K – \$70K	1	0	-
I-P20	1+ years of college (No degree)	Home-maker	Public/Rideshare	\$10K – \$20K	1	0	-
I-P21	Bachelor's degree	Self-employed	Car/Truck	\$60K – \$70K	2	0	Spouse/Partner

Table 8: Demographic information of interview participants.

Highest Degree		Employment		Commute Mode		Income		# Adults in Home		# Children in Home		Housemate(s)	
No schooling completed	0%	Full-time	37%	Bicycle	3%	<\$10K	10%	1	19%	0	75%	No one	16%
Nursery school	0%	Part-time	14%	Car/Truck	58%	\$10K – \$20K	8%	2	41%	1	12%	Roommate(s)	13%
Grades 1 through 11	0%	Unemployed	15%	Carpool/Vanpool	2%	\$20K – \$30K	9%	3	24%	2	8%	Spouse/Partner	35%
12th grade—no diploma	0%	Self-employed	12%	Motorcycle	1%	\$30K – \$40K	10%	4	14%	3	2%	Children	20%
Regular high school diploma	9%	Home-maker	4%	WFH	10%	\$40K – \$50K	10%	5	2%	4	3%	Parent(s)	32%
GED or alternative credential	1%	Student	15%	Public	10%	\$50K – \$60K	10%	>5	0%	5	0%	Other	14%
Some college credit	9%	Retired	3%	Rideshare	2%	\$60K – \$70K	8%			>5	0%		
1+ years of college (No degree)	20%			Walking	9%	\$70K – \$80K	6%						
Associate's degree	8%			Other	2%	\$80K – \$90K	6%						
Bachelor's degree	35%			-	9%	\$90K – \$100K	4%						
Master's degree	16%					\$100K – \$150K	13%						
Professional degree	1%					>\$150K	6%						
Doctorate degree	1%												

Table 9: Demographic information of survey participants. Note that the percentages under "Commute Mode" and "Housemate(s)" columns do not sum up to 100% as participants were able to choose any number of options.

F PERCEIVED BENEFITS AND HARMS IN THE PRESENTED SCENARIOS

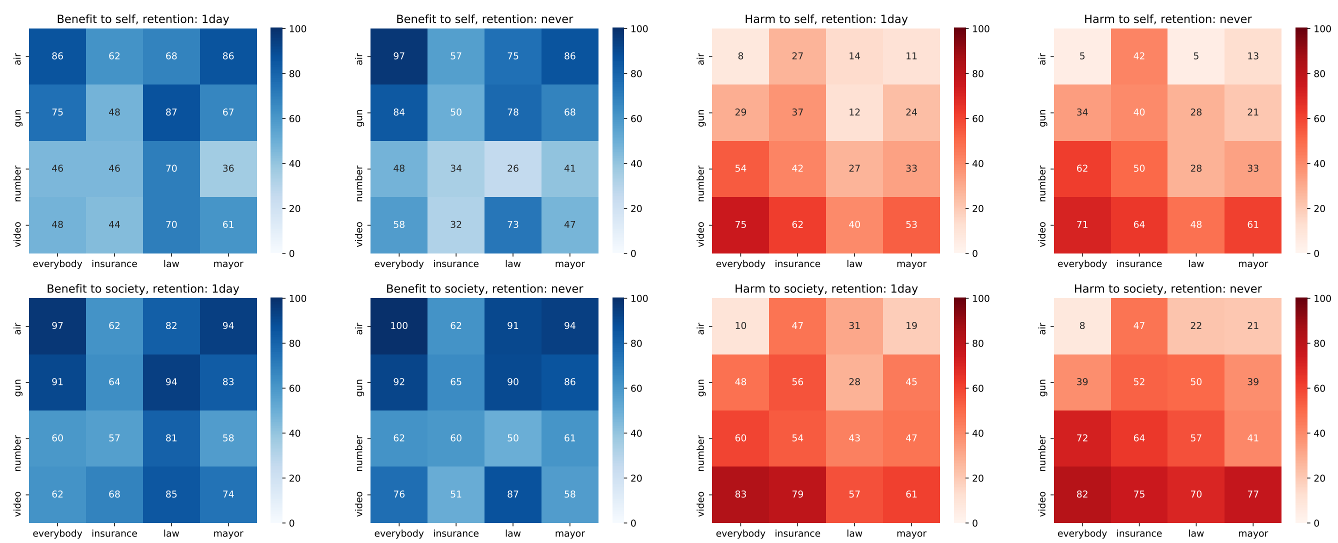

Figure 2: Percentages of perceived benefits and harms for various data collection and use scenarios.