Proceedings of the Nineteenth International AAAI Conference on Web and Social Media (ICWSM 2025)

"I'm Not for Sale" – Perceptions and Limited Awareness of Privacy Risks by Digital Natives About Location Data

Antoine Boutet,[1,2] **Victor Morel** [3,4]
[1] Insa-Lyon
[2] Inria
[3] Chalmers University of Technology
[4] University of Gothenburg
antoine.boutet@insa-lyon.fr
morelv@chalmers.se

Abstract

Although mobile devices benefit users in their daily lives in numerous ways, they also raise several privacy concerns. For instance, they can reveal sensitive information that can be inferred from location data. This location data is shared through service providers as well as mobile applications. Understanding how and with whom users share their location data as well as users' perception of the underlying privacy risks, are important notions to grasp in order to design usable privacy-enhancing technologies. In this work, we perform a quantitative and qualitative analysis of smartphone users' awareness, perception and self-reported behavior towards location data-sharing through a survey of $n = 99$ young adult participants (i.e., digital natives). We compare stated practices with actual behaviors to better understand their mental models, and survey participants' understanding of privacy risks before and after the inspection of location traces and the information that can be inferred therefrom.

Our empirical results show that participants have risky privacy practices: about 54% of participants underestimate the number of mobile applications to which they have granted access to their data, and 33% forget or do not think of revoking access to their data. Furthermore, most of the participants do not have a realistic perception of privacy risks and have generally heard little about privacy-related scandals. Also, by using a demonstrator to perform inferences from location data, we observe that slightly more than half of participants (57%) are surprised by the extent of potentially inferred information, and that 47% intend to reduce access to their data via permissions as a result of using the demonstrator. Last, a majority of participants have little knowledge of the tools to better protect themselves, but are nonetheless willing to follow suggestions to improve privacy (51%). Educating people, including digital natives, about privacy risks through transparency tools seems a promising approach.

1 Introduction

Smartphones have become the most popular electronic devices, used by 95% of all internet users[1]. The widespread adoption of these mobile devices with geolocation capabilities makes users' location traces[2] widely accessible to mobile services. Location is the most collected personal data by mobile applications (Achara et al. 2013), and while the exploitation of this data presents an obvious benefit to users in their daily lives – through personalized services, and potentially to the society with crowdsourcing initiatives (Stevens and D'Hondt 2010) –, an uncontrolled exploitation of this data continue to raise privacy concerns.

This location data can indeed reveal a lot of sensitive information about users and represents unique privacy risks and implications compared to other types of data shared via smartphones. For example, it has been shown that it is possible to infer personality traits, religious affiliation, sexual orientation, or health status (Gambs, Killijian, and del Prado Cortez 2010). Additionally, the uniqueness of location data provides a sort of fingerprint specific to each individual that can be used to re-identify users in anonymized data, i.e., little *a priori* knowledge about a user can be exploited to discriminate them in a large set of anonymous location data. For instance, journalists were able to re-identify and track the whereabouts of the former US president Trump from a large dataset (Stuart A. and Charlie 2019), and anonymous data was used to pinpoint Muslim cab drivers (Lorenzo 2015). The attack surface of the exploitation of personal data is not fully discovered. Regularly, new threats appear and are maliciously exploited for influential campaigns and interfering in political election[3], or discovering sensitive information[4].

With this increasing exposure of privacy risks, understanding how young users share their data, and the extent to which they are aware of security and privacy risks, are important properties to assess these risks and to develop effective Privacy and Transparency Enhancing Technologies matching current users' expectations to increase adoption. Although user perceptions of technology (Beckwith 2003) and the privacy paradox[5] (Barth et al. 2019; Kang and Jung 2021) have received a lot of attention, users' self-reported behaviors in mobility contexts associated with smartphones,

[1]https://datareportal.com/global-digital-overview

[2]That is, location data collected by Google from smartphones which gives more accurate and reliable information than other kinds of location data collection

[3]https://en.wikipedia.org/wiki/Troll_farm

[4]https://www.nytimes.com/2018/01/29/world/middleeast/strava-heat-map.html

[5]The privacy paradox refers to self-reported concerns about privacy appear to be in contradiction with often careless online behaviors.

and the impact of an awareness demonstration platform have been less studied in the academic literature. To fill this gap, this work addresses two research questions:

- RQ1: What are the perceptions and the understanding of young users' privacy and its protection in a mobility context?
- RQ2: What is the impact of a demonstrator for the visualization of location traces and associated privacy risks on these perceptions and understanding?

In order to answer our research questions, we designed and deployed survey questionnaires answered by $n = 99$ young participants (i.e., digital natives, persons who grew up in the information age, aged between 20 and 26 with an average of 21 in our case). Specifically, we explored the participants' behavior, self-reported behavior, and awareness regarding their own data-sharing practices. We devised a first questionnaire to study their perception of privacy and to their permission management with respect to location, in which we emphasized the eventual discrepancies between their remembered practices and their actual behaviors. We also surveyed participants' understanding of privacy risks before and after exposing them to location traces demonstrating what information can be inferred from this data, as well as their awareness of protection tools, in conjunction with a second questionnaire.

Our results show that participants have **risky practices in terms of privacy** where more than half of participants underestimate the number of mobile applications to which they have granted access to their data. In addition, most of the participants tend to forget about disabling location access permission for apps that are not actively used. Our results also show that participants are **poorly aware of the privacy risks** and are unable to list cases of personal data leaks or scandals linked to their uses despite the media coverage. Moreover, by using a demonstrator to perform inferences from location data, more than half of the participants (57%) are surprised by the extent of potentially inferred information and 47% intend to reduce access to their data via permissions. Finally, most of the participants are inclined to **better use protection tools in the future**, even if they are still little aware of the available tools to improve privacy today. Since our study focused mainly on digital natives, the transferability of the results to the rest of the population remains an open question. However, the academic knowledge in computer science of our participants certainly gives them a better perception of the risk to privacy than among other age groups. Our contributions are the following:

1. We provide the first insights on privacy perceptions in a mobility context associated with smartphones and mobile applications;
2. We study the impact of a demonstration platform of privacy risks visualization on privacy perceptions;
3. We formulate a set of recommendations to improve the management of privacy permissions on mobile.

The paper is organized as follows. Section 2 reviews the related work. Section 3 describes the considered methodology while the results are presented in Section 4. We discuss recommendations and highlight our limitations in Section 5, and conclude in Section 6.

2 Related Work

Prior research widely explored location privacy (Section 2.1) and users' perceptions of privacy (Section 2.2). However, little attention has been devoted to self-reported behaviors towards location data sharing. Moreover, to the best of our knowledge, no studies have used a demonstration platform to improve risk understanding in a way that would change users' perceptions about the balance between the benefits of invasive technologies and potential risks.

2.1 Privacy and Location

The privacy issues raised by location data gained a lot of traction in the last decade (Primault et al. 2019). In particular, user location traces extracted from various data have been shown to be highly unique (De Montjoye et al. 2013; Zang and Bolot 2011). This high uniqueness may act as a digital fingerprint and lead to the re-identification of users if their traces are associated with external knowledge. This uniqueness does not only concern location traces but characterizes all traces generated by a human (Andreas et al. 2016; Eckersley 2010; Rebekah and Rachel 2016), and can generate a risk of re-identification. Several cases of re-identification have been documented, for instance the re-identification of individuals from web search queries (Barbaro and Zeller 2006), taxi logs (Zhang and Wang 2016), or the notorious case of Governor William Weld using medical information (Barth-Jones 2012).

The uniqueness and risk of re-identification is not the unique threat related to location traces. Recent works have demonstrated that location is a very rich contextual information and leads to a strong inferential potential in terms of information that can be learned about individuals (**?**). For instance, location traces can reveal Points Of Interest (POI) of users such as their home and workplaces (Gambs, Killijian, and del Prado Cortez 2010), their race and gender (Zhong et al. 2015), their social network (Sharad and Danezis 2014), it can be used to predict their location patterns (Sadilek and Krumm 2012), to link accounts of the same user across different datasets (Riederer et al. 2016a), and infer even more sensitive information such as their religion or personality traits[6].

Due to the large adoption of mobile devices, the location data of users is extensively collected and shared (Almuhimedi et al. 2015; Englehardt and Narayanan 2016). The uncontrolled usage of this information can have an important impact on users such as unfair price discrimination (Mikians et al. 2012). The increasing use of connected devices collecting personal data and the lack of transparency on how the data is actually exploited raise privacy concerns. Only a handful of tools have been proposed to improve user awareness about the potential risk of revealing their location. For instance, (Riederer et al. 2016b) propose a tool to inspect

[6]Note that the processing of such sensitive information is prohibited by Article 9(1) of the General Data Protection Regulation (GDPR) on the processing of special categories of personal data.

the potential of location data, while (?) show the impact of protection mechanisms of the inference capabilities using a demonstration platform.

2.2 Perceptions of Privacy and Privacy Controls

Giving users the benefits of location services on their mobile devices while preserving their privacy is an ongoing challenge, as evidenced by mobile OSs iterating on the user interfaces for location notices and control every few years. For example, iOS allows users to select whether apps get fine- or coarse-grained location data[7], and will present popups when apps continually access location services in the background. Similarly, Android apps now have to request background location access separately from general location usage[8], and the OS will automatically revoke unused permissions from apps[9]. Even though mobile operating systems regularly improve user interfaces, the opaque privacy controls of location services still face criticism[10]. (Balash et al. 2022) explore the users' perceptions regarding access to Google accounts by mobile applications and formulate design recommendations to improve the current third-party management tools offered by Google, such as tracking recent access, automatically revoking access due to app disuse, and providing permission controls. (Wijesekera et al. 2015) analyze authorization preferences in different usage contexts and suggest determining the situations in which users would like to be confronted with security decisions.

Users' perceptions of the risks and benefits of technologies can determine their willingness to adopt them (Poikela and Kaiser 2016). More specifically, people are more likely to accept potentially invasive technology if they think its benefits will outweigh its potential risks (Beckwith 2003). Due to the massive adoption of location-enabled mobile applications, this fact suggests that users' perception of this trade-off is more in favor of the benefits than the potential risks. This attitude however depends on the perception of the said privacy risks. Studies have shown that users of mobile phones are often unaware of the data collected by apps running on their devices, and that a majority of users restrict some of their permissions (Almuhimedi et al. 2015) following a better awareness of data collection. Other studies explored the privacy paradox (Barth et al. 2019; Kang and Jung 2021) where self-reported concerns about privacy appear to be in contradiction with often careless online behaviors. However, another study focuses on university community (Gamarra et al. 2019) and show that this population of users does not have a genuine concern regarding the privacy of their geolocation data. Note that this paper did not study location data in mobility contexts associated with smartphones as we do.

Analyzing behavior and understanding users' perceptions are also important notions to grasp in order to further design Privacy and Transparency Enhancing Technologies (PETs and TETs). For instance, (Kaushik et al. 2021) conducted an online survey to understand people's perspectives on solely automated decision-making. They then formulate recommendations on how to design such systems. In a similar line of work, (Islami, Fischer-Hübner, and Papadimitratos 2022) conducted in-depth semi-structured interviews with 17 Swedish drivers to analyse their privacy perceptions and preferences for intelligent transportation systems, then to provide recommendations for suitable predefined privacy options. Several other works (Zufferey et al. 2023; Velykoivanenko et al. 2021; Mink et al. 2022) studied the perceptions of privacy and utility of users related to fitness-trackers and demonstrated a high potential for data minimization (i.e., reducing the volume of data sent to service provider). (Debatin et al. 2009), in turn, investigated Facebook users' awareness of privacy issues and perceived benefits and risks of utilizing Facebook and recommended better privacy protection, higher transparency and more education about the risks of posting personal information to reduce risky behavior. (Bielova et al. 2024) analysed the behaviors of websites' visitors through a study of the impact of dark patterns on consent decisions. (Veys et al. 2021) explored whether current data downloads (such as Google Takeout) actually achieve the transparency goals embodied by the right of access. Most participants indicated that current offerings need improvement to be useful, emphasizing the need for better filtration, visualization, and summarizing to help them hone in on key information. However, none of these related work specifically targets users' behavior and perception about data-sharing of location data. Although (Martin and Nissenbaum 2019) clearly addressed users' perceptions of location data, it is to be noted that it does so from an information science standpoint, closer to sociology and not from a UX/usability one. It nonetheless offers a relevant account for the interested reader.

Related to location data, (Farke et al. 2021) specifically analysed the user perceptions and reactions to Google's My Activity and how this web history dashboard increases or decreases end-users' concerns and benefits regarding data collection. Their results show that participants were surprised by the volume and detail of the collected data, but most of them were significantly more likely to be both 1) less concerned about data collection and 2) to view data collection more beneficially. However, this dashboard does not present any risks such as possible sensitive inferences associated with location traces or places visited. By also presenting the risks, our study shows that users are more concerned about privacy issues.

3 Methodology

We conducted two questionnaires on $n = 99$ young participants aged between 20 and 26 (with an average around 21) to collect quantitative and qualitative data about the participants' perception and self-reported behavior regarding the sharing of location data.

In a nutshell, the first questionnaire (hereinafter "**behavior questionnaire**") aimed at assessing their

[7]https://support.apple.com/guide/iphone/control-the-location-information-you-share-iph3dd5f9be/ios

[8]https://developer.android.com/about/versions/11/privacy/location

[9]https://developer.android.com/about/versions/11/privacy/permissions

[10]https://www.attorneygeneral.gov/taking-action/attorney-general-josh-shapiro-announces-391-million-settlement-with-google-over-location-tracking-practices/

self-reported practices, perceptions, and pre-conceived ideas about privacy. Following this first questionnaire, participants were then invited to use a demonstrator (denoted "**demonstration**" in what follows) which analyses location traces and reports several inferences as well as the impact of a protection mechanism using differential privacy (i.e., geo-indistiguishability (André s et al. 2013)). The second questionnaire (hereinafter "**perception questionnaire**") addressed their assessment of the risks.

3.1 Recruitment

We recruited young participants for our survey via a course offered to engineering students at [anonymized]. This process of recruitment ensured a high number of participants,[11] all of them using their smartphones. The participants were not compensated. The participants had followed computer science courses including data science, AI, and introduction to the GDPR. By mainly recruiting young computer engineers with knowledge of technologies, we performed our study on a homogeneous population which consumes digital information quickly, and which are supposedly aware of privacy risks, hence providing an "upper bound" of privacy perceptions in our context.

The study was performed over a four hours slot, during which students were successively invited to answer the behavior questionnaire, analyse privacy risks through the demonstration, and answer the perception questionnaire. Although the teacher of the course is one of the authors of the current paper, participants were presented with a consent form at the beginning of the survey, and they were able to decline without penalty (specified orally); only the experimentation of the demonstration platform was mandatory as part of their curriculum. The participation was anonymous, voluntary, and the course did not include any grade (one participant actually refused to answer the questionnaire, without penalty).

3.2 Ethical Considerations

The survey raises ethical issues in two main aspects.

First, the demonstration platform uses personal data, specifically location traces. This data is not necessarily sensitive in itself, but can act as a proxy for sensitive data. To reduce risks, participants used traces from data of the authors of this study. The data used during the demonstration phase was encrypted server-side and was deleted once all data had been processed (calculating Points of Interests, collecting metadata, and calculating inferences). The DPO of the authors' institution has validated the demonstration platform.

Second, we elicited answers from participants, which is also personal data. We took as many precautions as needed to guarantee 1) informed consent (see (Boutet and Morel 2025) for the consent form, and as noted in Section 3.1 they had the possibility to decline without any penalty), and 2) security of data storage and processing (authors only communicated using end-to-end encrypted tools, and the survey was conducted on an EU-based online tool stamped GDPR-compliant). One participant refused to answer the questionnaires.

Both the demonstration platform (regarding its compliance with the GDPR and other data security regulations) and the questionnaires were approved by the IRB of the academic institution both hosting the survey and where the students were based.

3.3 Design of the Questionnaires

We shortly describe in this section the organization of the two questionnaires. The interested reader will find their full content in (Boutet and Morel 2025). Each section is succinctly described in what follows:

- **Introduction:** This section presents the study and asks the participants for their consent to participate in the study and to collect their answers for research purposes.
- **Data-Sharing:** This section inquires about the participants' attitudes about data sharing (e.g., which apps they think could access their location data, with whom, etc).
- **Inferences:** This section questions their pre-conceived ideas about inferences from location data (e.g., *from* which data could their location be inferred, or whether the location is sufficient to single them out in a dataset).
- **Authorisation and Control:** This section touches upon authorisation and control of permissions in a mobile context. Questions in this section are about their revoking or giving access permissions to apps, or their evaluation of the difficulty to effectuate this permission management.
- **Privacy Concerns:** This section simply performs the analysis of an Internet User's Information Privacy Concerns (IUIPC) on a 7 Likert scale.
- **Demographics:** This section elicits demographics data.

The perception questionnaire is much shorter and is composed of questions regarding the possible (malicious) inferences conducted on their location traces, the protection against these privacy violations, and their new expectations and suggestions following the experimentation with the demonstration platform.

3.4 Risks Analysis - Demonstration

After the first behavior questionnaire, participants had to explore a demonstration platform informing them about privacy risks linked to location traces through example data. We use the (Boutet and Gambs 2019) demonstration platform. Participants were asked to use the platform to explore these location traces and to inspect the information that can be inferred. The demonstration typically offers to visualize one's traces per day (see Figure 1), which can raise awareness about how easy it is to identify Points of Interests (POIs) – such as the home, the workplace or any other visited places during the day –, and presents metadata (e.g., address, category of the place, attendance statistics).

The demonstration also offers a defense mechanism based on Differential Privacy (Dwork 2006) in the form of a slider, which controls the level of noise injected into the location data (see Figure 2). This feature is presented as a means to

[11](We address the limitations of our recruitment in Section 5.2).

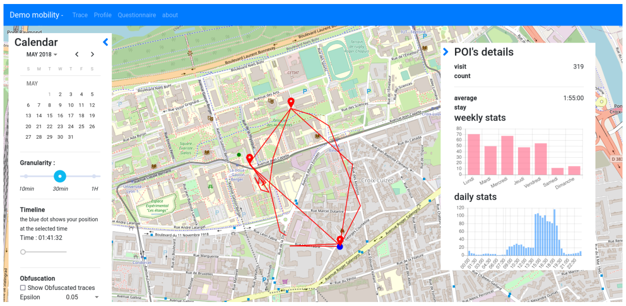

Figure 1: Screenshot of the demonstration's interface. Clicking on a POI yields attendance statistics.

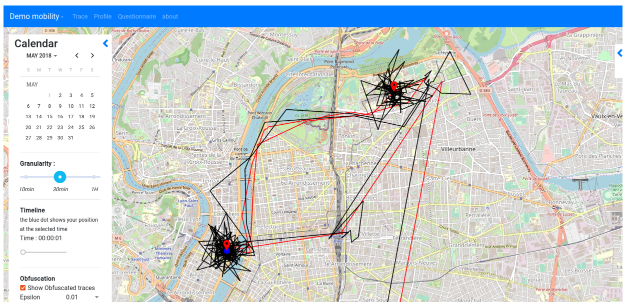

Figure 2: Illustration of the defense mechanism. Here the red lines show the initial traces, and the black lines represent the traces after the application of the noise.

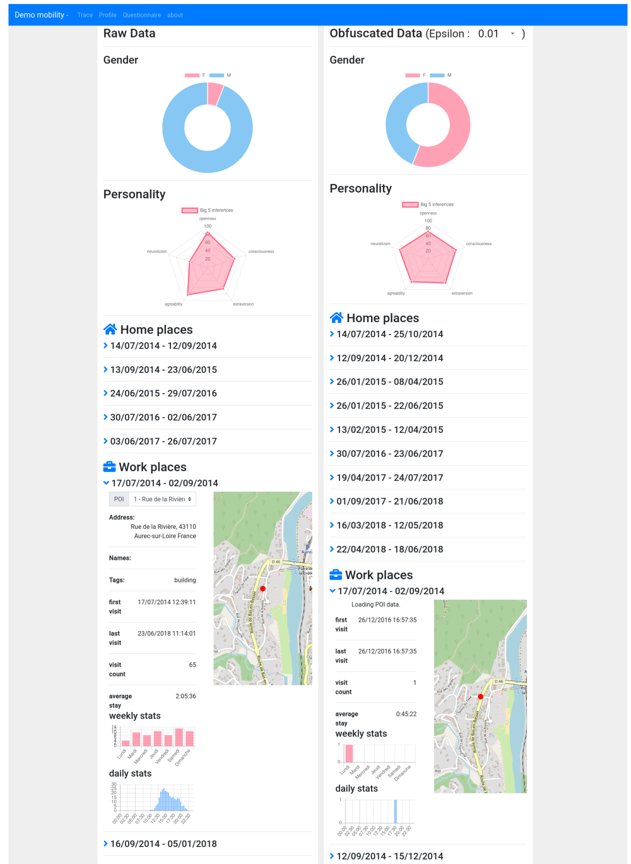

Figure 3: Inference performed by the platform. The left-hand side shows inference from raw data while the right-hand side presents the reduced risks after the application of the defense mechanism (i.e., geo-indistinguishability (André s et al. 2013)). The noisy version hardly yields the gender, smooths the personality traits, makes the daily and weekly attendance statistics unusable, and infers inaccurate POIs.

stimulate interest in privacy protection, and to raise awareness of the degradation of data quality when we increase protection with greater noise (to showcase the utility and privacy trade-off).

Finally, the demonstration highlights the risks raised by analyzing location traces and inferring gender, big five personality traits, home, and workplace (see Figure 3). Given the large number of inferences of sensitive information from location data reported in the literature (e.g., people encountered, religious affiliation, sexual orientation, or health status (Gambs, Killijian, and del Prado Cortez 2010)), only the presentation of these inferences fits fairly well into a representative risk of targeting (i.e., based on predictions computed on data collected and potentially exchanged between different parties).

3.5 Data Analysis

We performed a thematic analysis on the free text answered in the study (available in (Boutet and Morel 2025)) to analyse *qualitative* data. We took inspiration from the steps described in (Braun and Clarke 2006):

1. read the answers;
2. code the sentences from the previous step;
3. merge the codes in normalised codes for data analysis;
4. regroup the normalised codes from the previous step in themes;
5. review the themes and the related sentences to verify the homogeneity between them.

The analysis was performed by two independent annotators, and a consensus was reached at the end of the process. For most fields, a full thematic analysis was neither required nor necessarily desirable; curation and normalisation of the data was however necessary. Indeed, stopping at the third step was enough for several fields, such as the reasons given to revoke location permission. A full thematic analysis was performed on the answers to *Q32* "How do you think location data could be misused by a malicious person?", the themes of which were loosely derived from (Citron and Solove 2021) (Physical harms, Economic harms, Reputational harms, Informational integrity, Psychological harms), see (Boutet and Morel 2025).

We regrouped our codes for *Q36* "What are the measures imposed by the GDPR that reduce these risks?" under general legal categories (Rights, Consent, Clear purpose and transparency, Misc.) for statistical purposes, see (Boutet and Morel 2025). The rationale behind Q36 was to evaluate participants' knowledge about the GDPR.

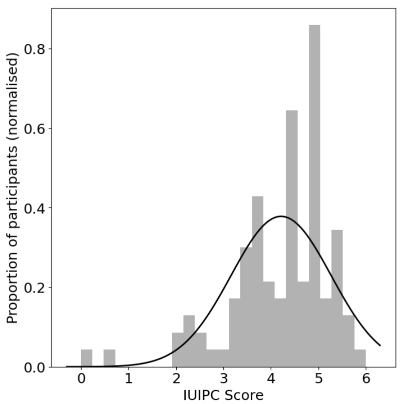

Figure 4: Plot of the Internet User's Information Privacy Concerns (IUIPC) distribution of participants.

3.6 General Statistics

Our participants are mostly equipped with Android phones: 60% against 40% for iOS users (*Q2*). Our population is not gender balanced, our sample was composed of 25% of women, 72% of men, and 3% prefer not to answer (*Q26*). Participants are young with an average age around 21 with a small standard deviation ($\sigma = 1.04$) (*Q27*).

Their privacy concerns have been assessed using the Internet Users' Information Privacy Concerns (IUIPC) Score. The IUIPC is a scale widely used in privacy research, it reflects Internet users' concerns about information privacy with a focus on "individuals' perceptions of fairness/justice in the context of information privacy". The distribution of IUIPC is depicted in Figure 4.

Note that our population is within average with respect to their privacy concerns (μ = 4.2, σ = 1.05)[12] in spite of young adults being generally less concerned about their privacy than older populations (as mentioned in Section 2).

4 Results

This section showcases the most relevant results for our research questions, with the first two subsections dedicated to answering RQ1 ("What are the perceptions and the understanding of young users' privacy and its protection in a mobility context?"), and the last subsection more tailored to RQ2 ("What is the impact of visualization of location traces and associated privacy risks on these concepts?").

4.1 Privacy Risky Practices

Participants do not adopt safe practices in terms of privacy. Specifically, they tend to underestimate the number of apps that have access to the location, and they tend not to think about disabling location access for apps that are not actively or forget to do so.

Underestimation of the number of apps that have access to location. Participants were asked to estimate how many mobile apps have access to the location on their smartphone "on top of their head" (*Q3*), and this data was then compared to the actual number (both for "always" and for "when using the app") (*Q8*). Unsurprisingly, the majority of participants (52) underestimated the number of apps with location access, 7 overestimated it, but 38 were correct (with 17 on the open range, i.e., they selected the maximum value), as presented in Figure 5. We interpret this result as a wrongful mental model due to 1) the lack of transparency of apps regarding their data collection, 2) an ever-increasing number of applications, combined with 3) a lack of a central controlling system for privacy permissions.

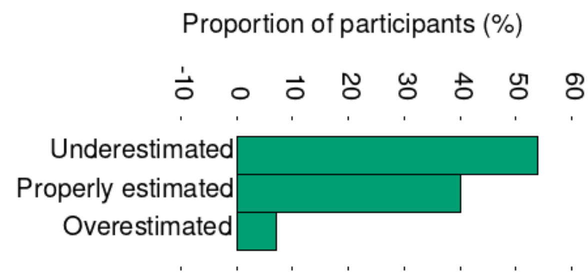

Figure 5: Participants tend to underestimate the number of apps that have access to the location (*Q3* against *Q8*).

While the average number of apps with continuous access to the location was pretty low but with a high standard deviation (μ = 2.74, σ = 6.32) – meaning that most participants do not allow continuous access to their location except for a few outliers with a high number of apps –, we note that a much higher number of apps had access to location when on use (μ = 11.04, σ = 8.93)(*Q8*).

Amongst the applications with access to location (*Q9*), Google Maps is leading the list "on top of their heads" (38 mentions), but has more access in practice (15 for "always" and 27 "when using the app"); it is followed by Instagram (respectively 27, 4, 31) and Snapchat (26, 4, 22).

Participants tend to forget about data-sharing. The results show that most participants had installed numerous applications with access to location (*Q8*). In addition, we see that 74.2% (46 actively use most, 26 only some of them) have at least one application not actively used but with current access to location data (*Q10*)[13]. When asked why these unused applications still have access to their location (*Q11*), a majority of the participants have either forgotten (11) or not thought (12) about removing access (see Figure 6). Another proportion of participants are either lazy (6) or unaware (7) that these applications collect location data. Finally, the rest of the participants (15) prefer to leave the permission to the app in case they need to use it.

[12]See (Zufferey et al. 2023), for an example of less concerned participants (μ = 3.5, σ = 1.6) and (Naeini et al. 2017) for more concerned participants with μ between 4.79 and 5.44 normalized to a [0-6] scale.

[13]We orally clarified to participants that "actively used" amounts to a weekly use or more frequent.

Basically, participants have difficulties remembering their apps' privacy practices by themselves, which make the latest features of iOS ("Offload Unused Apps") and Android ("Unused Apps") all the more relevant.

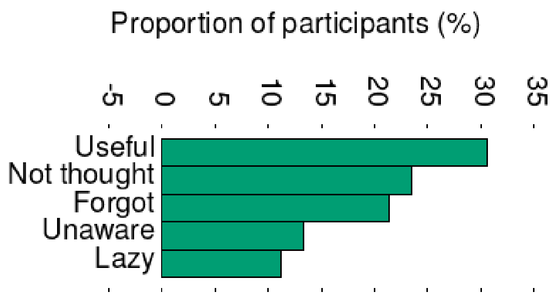

Figure 6: Participants tend not to think about disabling location access for apps that are not actively used or forget to do so (*Q11*).

Authorisation and Control. We asked participants questions regarding the control of location access to mobile applications (*Q23*). Only 4% think it is not possible to revoke (cancel) previously granted mobile app access (*Q18*) and 16% of the participants never modified the privacy settings related to their location data (*Q19*). A greater proportion of participants reduced the availability of their location data than those who increased their availability (75% versus 8%). For those who said they had already revoked access, the reasons given by 68% of participants felt that this access was not necessary for the application ("I'm not for sale"); or did not understand why this access was necessary (*Q21*).

4.2 Participants are Poorly Aware of the Risks

They do not know how location can be captured, they are not aware of the great inference capacity associated with the analysis of location data, and they are unable to cite cases of personal data leaks or scandals related to their use despite increased media coverage.

Participants have misconceptions about how location can be captured. To assess the perception of privacy risks linked to the sharing of location information, we asked the participants how a mobile application could capture location (*Q12*). The responses are displayed in Figure 7 and show that not all participants are aware of the localization capabilities of the main technologies embedded in a phone. Although GPS is identified by almost all participants (96%), we see that only 85%, 64% and 78% of participants are aware that respectively WiFi, Bluetooth, and IP addresses can be used as proxies to infer location.

Unaware of the inference capacity from location data. In the behavior questionnaire before the demonstration, we then asked participants to identify what information could be inferred from location amongst a list of information generally categorised as sensitive (*Q13*)[14]. The results are represented in Figure 8 and show that the risks of inference are

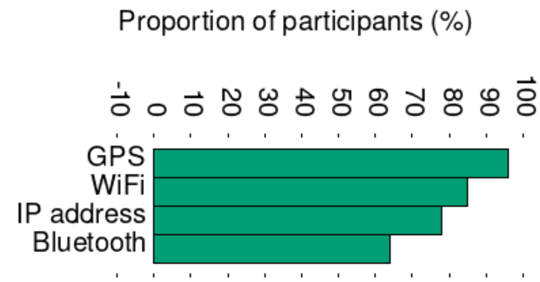

Figure 7: Proportion of participants aware of source of data from which location can be inferred (*Q12*).

moderately known. Although all participants perceived that the place of work can be inferred from location traces, only 94% think that the place of residence can also be inferred. These inferences are based on the time spent in a location over office hours and overnight stays. This change in perception raises questions about the rationality or basis for the response because we spend more time in the home than in the workplace. Finally, we observe that 10% of participants do not believe that this information can lead to re-identification (*Q15*), which is a surprisingly high number considering that all participants had previously followed courses on data science and AI. On the other hand, when asked why a service provider might be interested in the location of individuals (*Q14*), 88% responded that this information is for marketing purposes. So they still have the perception that this information could be used for targeting or profiling.

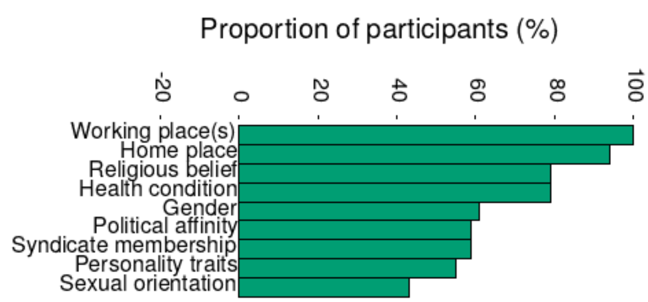

Figure 8: Answers to *Q13* "Do you think it is possible to infer the following information from the location?".

Little echo about data leaks or privacy scandals. The majority of participants are unable to list cases of personal data leaks or scandals linked to their uses despite press articles regularly published on the subject (*Q33*). In view of the numerous scandals and cases linked to personal data reported very regularly by the press (e.g., ransomware and data leaks in hospitals, attempted influence on social networks), this lack of awareness of the risks raises questions about their exposure to the media. Only 32 knew about a scandal, half of these cases were Cambridge Analytica (16 mentions), while major scandals such as Pegasus or the Snowden leaks were barely mentioned (3 mentions each). The medium-high reported IUIPC of the participants (see Section 3.6) does not

[14] Most of these categories are listed as special categories of data by Article 9(1) of the GDPR, see Section 2.1.

seem to come from exposure to cases concerning privacy violations reported by the press.

4.3 The risk Demonstrator Increased Awareness of Participants About Privacy

After having experimented with the demonstration, a third of the participants (33) were planning on limiting the access to mobile privacy permissions, and deleting the history of mobility on the Google Takeout platform (*Q37*). We contend that the visualisation of location traces and the risks associated with the inference increased awareness of participants towards privacy, and encouraged them to adopt more privacy-preserving behaviors.

Relative surprise for the number of possible inferences. Once the participants have been able to analyze location traces and discover the information that can be learned with the demonstrator, more than half of the participants (56%) declared being surprised by the number of inferences made possible through location data (*Q29*). Additionally, when asked whether they feel well protected against the various threats observed, a large majority of participants responded that they felt unprotected (84%) (*Q35*). Knowing this inference ability, participants were asked if they think Google analyzes users' location data (*Q30*), and if they think Google shares this data with third parties (*Q31*). The results show that 93% of participants believed that Google analyzes their location data and 89% believed that their data is shared with third parties. Even after the inspection of location traces with the platform, 7% of participants still did not think that location data can be used to re-identify users (*Q34*).

Worries about surveillance, economic, and physical harms. The most worrisome inferences are informational harms (surveillance, stalking[15], inference, data selling, targeted advertising; 88 mentions), followed by economic harms (blackmail, robbery, scam, identity theft; 51 mentions), and finally physical harms (physical harm, harassment; 31 mentions) (*Q32*).

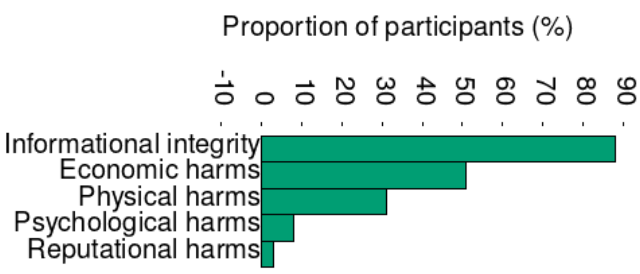

Figure 9: The most worrisome inferences are informational harms (i.e., surveillance, stalking, inference, data selling, targeted advertising) (*Q32*).

We observed interesting preconceived ideas regarding those harms. First, in spite of being in the most cited category, targeted advertising was specifically mentioned only three times, whereas it is the main privacy-invasive goal of location data analysis by Google (from which the traces were extracted). Second, robbery was over-represented in the answers (28 mentions, 2nd most cited code) although it does not happen so often in practice[16].

Positive attitude towards protective mechanisms. Despite the lack of awareness about the risks, participants are inclined to use protection tools. When asked whether technologies could reduce and provide greater control over personal information disclosed via location data, 76.3% of participants responded that they would be willing to use Privacy Enhancing Technologies (PETs) (*Q16*). Consistently, 51.5% declared to plan using PETs after having experimented with the demonstration platform (*Q38*). However, a lower score on question Q38 than on question Q16 means that participants do not have a concrete idea of existing privacy-enhancing tools. On the other hand, 84% of participants said they have never looked to see if this type of technology or tools existed (*Q39*).

5 Discussion

This section formulates practical recommendations and highlights the limitations of our work.

5.1 Recommendations

We craft in this section recommendations for the design of better privacy-friendly systems, and of Transparency Enhancing Technologies (TETs) to foster privacy awareness. These recommendations are drawn from both participants' insights and our own interpretations.

Assist in the revocation of apps. Our results show that participants tend to forget about data-sharing, where more than a quarter of the participants declared not thinking about revoking access (12), forgot to do it (11), or are simply too lazy (6), and that they have inaccurate representations of the number of apps with access to their location (see Section 4.1). However, we note that 14 participants suggested including a system-wide reminder of which apps can access location as recommendations for better and more usable Privacy Enhancing Technologies (PETs). Note that the latest version of Android and iOS indicate when an app accesses location, and Android now offers a centralized interface indicating which apps require and use geolocation. This centralized interface can be hard to find amongst the different settings, and no participant were aware of its existence. However, both Android and iOS lack a centralized interface to inform which app is *currently* accessing location, nor do they provide fine-grain data regarding its use (i.e., how many times did each app request geolocation). We therefore recommend that **mobile OSs integrate a centralised and easily accessible interface about location usage**, providing not only which geolocation permission has been granted to each app but also a fine-grain data regarding its uses.

[15] Here, *stalking* refers to close surveillance without intervention, we coded surveillance with physical and/or oral intervention under *harassment*.

[16] A possible explanation can be that of the black swan theory, or the presence of websites raising awareness of this specific risk such as https://pleaserobme.com/

More transparency. Our results also show that knowing which apps have access to location is not enough; 16 participants also proposed more transparency and more specifically to be able to be informed at any time of which applications are collecting data. Seeing the lack of perception of privacy risks and the impact of tools allowing to calculate and visualize these risks on this perception, we suggest improving transparency by **providing users with statistics and metrics of risk** associated with data collection for each application. We believe that this personalized risk analysis would be an important lever to help users to evaluate the trade-off between utility and respect for their privacy and to better calibrate what they wish to share.

Access privacy permissions from the app. Another highlight from the participants' answers is the possibility to access privacy permissions directly from within the app. Indeed, today's privacy permissions (on both iOS and Android) are mainly available through the general settings of the smartphone, which can be hard to access for lay users[17]. We therefore recommend mobile OSs to **include a shortcut to an app's privacy permissions from the app itself**, for instance from the status bar or by executing a special gesture. This suggested feature is currently unavailable on modern mobile operating systems and has not been suggested anywhere else to the best of our knowledge.

5.2 Limitations

First, our demographics do not represent a fair sample of digital users. Our participants were mostly males, all of them young adults educated in a top European engineering school, although with different backgrounds and technical skills. This bias in our representativeness played a part in the socio-economic backgrounds of participants, who are most likely to come from well-off social environments. Recruiting participants in universities often faces this limitation in academic studies on usages. For instance, (Kulyk et al. 2022) involved 40 participants mostly young and well-educated, while (Karegar, Pettersson, and Fischer-Hübner 2020) involved 80 participants, mostly young, with 45% in the 18–25 age group and 20% between the ages of 26 and 35, a majority ($n = 54$) were bachelor students or had their bachelor degrees. (Almuhimedi et al. 2015) involved 23 participants (65% female; ages 18–44, median=23), and (Kaushik et al. 2021) involved 392 participants ranging from 25 to 35.

The transferability of the result to the rest of the population therefore remains an open question. However, given that our sample is predominantly composed of "digital natives" – who are often assumed to be more comfortable with personal information sharing and more knowledgeable about digital technologies –, we expected a better perception about the privacy risk than amongst other age groups. This expectation also has to be put in the perspective of their curriculum: students in computer engineering with academic knowledge of data science, AI, and rudiments of GDPR.

Second, the scope of our work must be restricted to a specific type of demonstrator, our results should not be generalized to all visualization and awareness tools. Moreover, the results largely presented self-reported and planned behaviors. Therefore, an interesting research avenue for future work is to study the *actual behavioral* changes introduced by the experiment of the tool, and long-term behavior analysis through *longitudinal studies*. Perceptions can also change according to the type of apps and their access to location data, as well as other sensitive personal data collected.

Third, participants used example data on the demonstration platform, which is perhaps less illustrative than if they had inspected their own location data. We envision working with personalized examples in the future. Our results also lack perspective on whether our participants were able to deeply understand the goal and the meaning of obfuscated traces and the epsilon number (see (Karegar, Alaqra, and Fischer-Hübner 2022)).

[17] Note that it is possible to access privacy permissions for an app from some Android launchers, but **not** from within the app.

6 Conclusion

Through a survey involving $n = 99$ young mobile phone users, this work contributes to a better understanding of the practices in relation to the sharing of their location data with mobile applications as well as the associated privacy issues. By qualitatively and quantitatively analyzing user perceptions and self-reported sharing behaviors, we provide valuable insights for researchers and privacy practitioners to better understand users and better design new Privacy Enhancing Technologies (PETs) and Transparency Enhancing Technologies (TETs) related to location data. We also showed the importance of displaying a more interactive risk analysis to make users better aware of privacy risks.

Acknowledgements

This work has been supported by the ANR 22-PECY-0002 IPOP (Interdisciplinary Project on Privacy) project of the Cybersecurity PEPR, and partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation. We would like to thank Farzaneh Karegar and Nataliia Bielova for their feedback on earlier drafts, and anonymous reviewers who helped us improve this article.

References

Achara, J. P.; Castelluccia, C.; Lefruit, J.-D.; Roca, V.; Baudot, F.; and Delcroix, G. 2013. Mobilitics: Analyzing Privacy Leaks in Smartphones. *ERCIM News*.

Almuhimedi, H.; Schaub, F.; Sadeh, N.; Adjerid, I.; Acquisti, A.; Gluck, J.; Cranor, L. F.; and Agarwal, Y. 2015. Your location has been shared 5,398 times! A field study on mobile app privacy nudging. In *CHI*, 787–796.

Andreas, K.; Hugo, G.; Tobias, B.; Konrad, R.; and Felix, F. 2016. Fingerprinting Mobile Devices Using Personalized Configurations. *PoPETs*, 2016(1): 4–19.

André s, M. E.; Bordenabe, N. E.; Chatzikokolakis, K.; and Palamidessi, C. 2013. Geo-indistinguishability. In *CCS*.

Balash, D. G.; Wu, X.; Grant, M.; Reyes, I.; and Aviv, A. J. 2022. Security and Privacy Perceptions of Third-Party Application Access for Google Accounts. In *USENIX Security*, 3397–3414.

Barbaro, M.; and Zeller, T. 2006. A Face is exposed for AOL searcher no. 4417749. *New York Times*.

Barth, S.; de Jong, M. D.; Junger, M.; Hartel, P. H.; and Roppelt, J. C. 2019. Putting the privacy paradox to the test: Online privacy and security behaviors among users with technical knowledge, privacy awareness, and financial resources. *Telematics and Informatics*, 41: 55–69.

Barth-Jones, D. 2012. The 'Re-Identification' of Governor William Weld's Medical Information: A Critical Re-Examination of Health Data Identification Risks and Privacy Protections, Then and Now. *SSRN Electronic Journal*.

Beckwith, R. 2003. Designing for ubiquity: the perception of privacy. *IEEE Pervasive Computing*, 2(2): 40–46.

Bielova, N.; Chammat, M.; Toubiana, V.; Hary, E.; Nguyen, A.; and Litvine, L. 2024. The Effect of Design Patterns on (Present and Future) Cookie Consent Decisions: Supplemental Materials. In *USENIX Security*.

Boutet, A.; and Gambs, S. 2019. Demo: Inspect what your location history reveals about you Raising user awareness on privacy threats associated with disclosing his location data. In *CIKM*, 2861–2864.

Boutet, A.; and Morel, V. 2025. "I'm not for sale" – Perceptions and limited awareness of privacy risks by digital natives about location data. arXiv:2502.11658.

Braun, V.; and Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2): 77–101.

Citron, D. K.; and Solove, D. J. 2021. Privacy Harms. *SSRN Electronic Journal*.

De Montjoye, Y.-A.; Hidalgo, C. A.; Verleysen, M.; and Blondel, V. D. 2013. Unique in the crowd: The privacy bounds of human mobility. *Scientific reports*, 3(1): 1–5.

Debatin, B.; Lovejoy, J. P.; Horn, A.-K.; and Hughes, B. N. 2009. Facebook and Online Privacy: Attitudes, Behaviors, and Unintended Consequences. *Journal of Computer-Mediated Communication*, 15(1): 83–108.

Dwork, C. 2006. Differential Privacy. In *Automata, Languages and Programming*, volume 4052, 1–12. Berlin, Heidelberg: Springer Berlin Heidelberg.

Eckersley, P. 2010. How unique is your web browser? In *PETS*, 1–18.

Englehardt, S.; and Narayanan, A. 2016. Online Tracking: A 1-Million-Site Measurement and Analysis. In *CCS*, 1388–1401.

Farke, F. M.; Balash, D. G.; Golla, M.; Dürmuth, M.; and Aviv, A. J. 2021. Are Privacy Dashboards Good for End Users? Evaluating User Perceptions and Reactions to Google's My Activity. In *USENIX Security*, 483–500.

FORCE11. 2020. The FAIR Data principles. https://force11.org/info/the-fair-data-principles/.

Gamarra, M.; Meriño Fuentes, I.; Calabria Sarmiento, J.; Gutierrez Acosta, O.; Barrios Barrios, M.; Leal, N.; and Wightman Rojas, P. M. 2019. Privacy Perception in Location-Based Services for Mobile Devices in the University Community of the North Coast of Colombia. 23(1).

Gambs, S.; Killijian, M.-O.; and del Prado Cortez, M. N. n. 2010. Show Me How You Move and I Will Tell You Who You Are. In *SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS*, 34–41.

Gebru, T.; Morgenstern, J.; Vecchione, B.; Vaughan, J. W.; Wallach, H.; Iii, H. D.; and Crawford, K. 2021. Datasheets for datasets. *Communications of the ACM*, 64(12): 86–92.

Islami, L.; Fischer-Hübner, S.; and Papadimitratos, P. 2022. Capturing drivers' privacy preferences for intelligent transportation systems: An intercultural perspective. *Computers & Security*, 123: 102913.

Kang, H.; and Jung, E. H. 2021. The smart wearables-privacy paradox: A cluster analysis of smartwatch users. *Behaviour & Information Technology*, 40(16): 1755–1768.

Karegar, F.; Alaqra, A. S.; and Fischer-Hübner, S. 2022. Exploring {User-Suitable} metaphors for differentially private data analyses. In *SOUPS*, 175–193.

Karegar, F.; Pettersson, J. S.; and Fischer-Hübner, S. 2020. The Dilemma of User Engagement in Privacy Notices: Effects of Interaction Modes and Habituation on User Attention. *ACM Trans. Priv. Secur.*, 23(1).

Kaushik, S.; Yao, Y.; Dewitte, P.; and Wang, Y. 2021. "How I Know For Sure": People's Perspectives on Solely Automated Decision-Making (SADM). In *SOUPS*, 159–180.

Kulyk, O.; Rafnsson, W.; Borberg, I.; and Pedersen, R. 2022. “So I Sold My Soul”: Effects of Dark Patterns in Cookie Notices on End-User Behavior and Perceptions. In *SOUPS*.

Lorenzo, F.-B. 2015. Redditor cracks anonymous data trove to pinpoint Muslim cab drivers.

Martin, K. E. M.; and Nissenbaum, H. F. 2019. What Is It About Location? *SSRN Electronic Journal*.

Mikians, J.; Gyarmati, L.; Erramilli, V.; and Laoutaris, N. 2012. Detecting Price and Search Discrimination on the Internet. In *Workshop on Hot Topics in Networks*, 79–84.

Mink, J.; Yuile, A. R.; Pal, U.; Aviv, A. J.; and Bates, A. 2022. Users Can Deduce Sensitive Locations Protected by Privacy Zones on Fitness Tracking Apps. In *CHI*.

Naeini, P. E.; Bhagavatula, S.; Habib, H.; Degeling, M.; Bauer, L.; Cranor, L.; and Sadeh, N. 2017. Privacy Expectations and Preferences in an IoT World. In *SOUPS*.

Poikela, M. E.; and Kaiser, F. 2016. ”It Is a Topic That Confuses Me” – Privacy Perceptions in Usage of Location-Based Applications. In *EuroUSEC*.

Primault, V.; Boutet, A.; Mokhtar, S. B.; and Brunie, L. 2019. The Long Road to Computational Location Privacy: A Survey. *IEEE Communications Surveys & Tutorials*, 21(3): 2772–2793.

Rebekah, O.; and Rachel, G. 2016. Blogs, Twitter Feeds, and Reddit Comments: Cross-domain Authorship Attribution. *PoPETS*, 2016(3): 155–171.

Riederer, C.; Kim, Y.; Chaintreau, A.; Korula, N.; and Lattanzi, S. 2016a. Linking Users Across Domains with Location Data: Theory and Validation. In *WWW*, 707–719.

Riederer, C. J.; Echickson, D.; Huang, S.; and Chaintreau, A. 2016b. FindYou: A Personal Location Privacy Auditing Tool. *WWW*.

Sadilek, A.; and Krumm, J. 2012. Far Out: Predicting Long-Term Human Mobility. In *AAAI*.

Sharad, K.; and Danezis, G. 2014. An Automated Social Graph De-Anonymization Technique. In *WPES*, 47–58.

Stevens, M.; and D'Hondt, E. 2010. Crowdsourcing of pollution data using smartphones. In *Workshop on ubiquitous crowdsourcing*, 1–4.

Stuart A., T.; and Charlie, W. 2019. How to Track President Trump.

Velykoivanenko, L.; Niksirat, K. S.; Zufferey, N.; Humbert, M.; Huguenin, K.; and Cherubini, M. 2021. Are Those Steps Worth Your Privacy?: Fitness-Tracker Users' Perceptions of Privacy and Utility. *IMWUT*, 5(4): 181:1–181:41.

Veys, S.; Serrano, D.; Stamos, M.; Herman, M.; Reitinger, N.; Mazurek, M. L.; and Ur, B. 2021. Pursuing Usable and Useful Data Downloads Under GDPR/CCPA Access Rights via Co-Design. In *SOUPS*, 217–242.

Wijesekera, P.; Baokar, A.; Hosseini, A.; Egelman, S.; Wagner, D.; and Beznosov, K. 2015. Android Permissions Remystified: A Field Study on Contextual Integrity. In *USENIX Security*, 499–514.

Zang, H.; and Bolot, J. 2011. Anonymization of Location Data Does Not Work: A Large-scale Measurement Study. In *MobiCom*, 145–156.

Zhang, S.; and Wang, Z. 2016. Inferring passenger denial behavior of taxi drivers from large-scale taxi traces. *PloS one*, 11(11): e0165597.

Zhong, Y.; Yuan, N. J.; Zhong, W.; Zhang, F.; and Xie, X. 2015. You Are Where You Go: Inferring Demographic Attributes from Location Check-Ins. In *WSDM*, 295–304.

Zufferey, N.; Niksirat, K. S.; Humbert, M.; and Huguenin, K. 2023. "Revoked just now!" Users' Behaviors Toward Fitness-Data Sharing with Third-Party Applications. *PETS*, 2023(1): 47–67.

Paper Checklist

1. For most authors...
   (a) Would answering this research question advance science without violating social contracts, such as violating privacy norms, perpetuating unfair profiling, exacerbating the socio-economic divide, or implying disrespect to societies or cultures? *Yes, see Section 3.2 on Ethical Considerations*
   (b) Do your main claims in the abstract and introduction accurately reflect the paper's contributions and scope? *Yes*
   (c) Do you clarify how the proposed methodological approach is appropriate for the claims made? *Yes*
   (d) Do you clarify what are possible artifacts in the data used, given population-specific distributions? *Yes, see Section 3.5*
   (e) Did you describe the limitations of your work? *Yes, in Section 5.2*
   (f) Did you discuss any potential negative societal impacts of your work? *NA*
   (g) Did you discuss any potential misuse of your work? *NA*
   (h) Did you describe steps taken to prevent or mitigate potential negative outcomes of the research, such as data and model documentation, data anonymization, responsible release, access control, and the reproducibility of findings? *Yes, in Section 3.2*
   (i) Have you read the ethics review guidelines and ensured that your paper conforms to them? *Yes*
2. Additionally, if your study involves hypotheses testing...
   (a) Did you clearly state the assumptions underlying all theoretical results? *NA*
   (b) Have you provided justifications for all theoretical results? *NA*
   (c) Did you discuss competing hypotheses or theories that might challenge or complement your theoretical results? *NA*
   (d) Have you considered alternative mechanisms or explanations that might account for the same outcomes observed in your study? *NA*
   (e) Did you address potential biases or limitations in your theoretical framework? *NA*
   (f) Have you related your theoretical results to the existing literature in social science? *NA*
   (g) Did you discuss the implications of your theoretical results for policy, practice, or further research in the social science domain? *NA*
3. Additionally, if you are including theoretical proofs...
   (a) Did you state the full set of assumptions of all theoretical results? *NA*
   (b) Did you include complete proofs of all theoretical results? *NA*
4. Additionally, if you ran machine learning experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? *NA*
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? *NA*
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? *NA*
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? *NA*
   (e) Do you justify how the proposed evaluation is sufficient and appropriate to the claims made? *NA*
   (f) Do you discuss what is "the cost" of misclassification and fault (in)tolerance? *NA*
5. Additionally, if you are using existing assets (e.g., code, data, models) or curating/releasing new assets, **without compromising anonymity**...
   (a) If your work uses existing assets, did you cite the creators? *Yes, in Section 3.4*
   (b) Did you mention the license of the assets? *NA*

(c) Did you include any new assets in the supplemental material or as a URL? *NA*

(d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? *Yes, in Section 3.2*

(e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? *Yes, in Section 3.2*

(f) If you are curating or releasing new datasets, did you discuss how you intend to make your datasets FAIR (see FORCE11 (2020))? *NA*

(g) If you are curating or releasing new datasets, did you create a Datasheet for the Dataset (see Gebru et al. (2021))? *NA*

6. Additionally, if you used crowdsourcing or conducted research with human subjects, **without compromising anonymity**...

(a) Did you include the full text of instructions given to participants and screenshots? *Yes, in appendix for the submission, and removed in the final version for space limitations, but available in (Boutet and Morel 2025)*

(b) Did you describe any potential participant risks, with mentions of Institutional Review Board (IRB) approvals? *Yes, see Section 3.2 on Ethical Considerations*

(c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? *Yes, in Section 3.1*

(d) Did you discuss how data is stored, shared, and deidentified? *Yes, in Section 3.2*