“It Is a Topic That Confuses Me” – Privacy Perceptions in Usage of Location-Based Applications

Maija E. Poikela, and Felix Kaiser
Quality and Usability Lab,
Technische Universität Berlin
Berlin, Germany
Email: maija.poikela@qu.tu-berlin.de

***Abstract*— Location-based applications bring ever more possibilities for the users: finding a soulmate, locating good restaurants in vicinity, or tracking a lost phone. Benefits are abundant, but, whether the user realizes it or not, so are the risks. This raises questions about what the users of location-based applications think happens with their location data, whether they see the usage as a tradeoff between benefits and risks, and whether they attempt to protect themselves from privacy risks. We conducted a set of semi-structured interviews (N = 41) with an explorative approach to investigate smartphone users' perceptions of location-based applications. Among other things, we investigated the benefits that have been experienced, the risks that cause concern, and the expectations of what happens with the location data. The data was then analyzed to further study the relationships between these concepts. Our results suggest that trusting individuals see more benefits in location-based applications than others, and on the other hand, those who express mistrust report more risks than others. Interestingly, participants with some limitations in their knowledge of location-based applications said more often than others that there are no risks in using location-based applications. On the other hand, participants with good knowledge seem to be protecting themselves from privacy risks more.**

***Index Terms*—knowledge, location, location-based applications, privacy, trust**

I. Introduction

In the age of information technology, the nature of interaction has changed. Unlike in physical world, in online social interactions the audience with whom one interacts is no more physically or temporarily restricted [1]. A comment posted in an online forum today might get a different kind of interpretation if re-posted and read in a different context, by an audience not originally imagined by the person. This kind of breaking of *privacy boundaries* cause discomfort and privacy issues, since the user is no longer in control of their data [2]. Similarly, when a user's personal information is used in a context not intended by the individual, boundary turbulence ensues. This applies also for location information.

Location can be considered personally identifiable information, as from one's movement patterns, a whole range of personal details can be inferred. If location data is also combined with other data such as medical data, or internet searches, a great deal can be inferred about an individual.

To protect oneself from privacy breaches and to be in control of one's personal information, one should be knowledgeable of what happens with the data. However understanding what happens with one's data when using online services is non-trivial, and in fact most users have been shown to have no understanding about the data flow, nor about its usage [3]. Privacy policies are lengthy [4], and written in a language incomprehensible to the common user [5].

The number of location-based applications (LBA) has drastically increased within the last decade as smartphones have gained popularity. The location of a device can be calculated using one or several methods, including triangulation based on cell towers, satellites, or Wi-Fi signals. Using more than one of the methods improves the accuracy of the location and overcomes issues in some methods (e.g. the satellite signal getting affected by blocking objects, such as buildings). This information is available for applications installed on the users' devices and is retrieved either at certain intervals (for example every 5 minutes), or when requested by the user. Location-based applications use this geographical location of a device, providing mobile users a number of functionalities. These include services that use location for finding information such as nearby restaurants, locating one's lost device, or for social purposes, including finding a partner, or enhancing one's social status through location check-ins. The location information can also be saved to the user's profile and the movement traces can be used to provide

EuroUSEC '16, 18 July 2016, Darmstadt, Germany

http://dx.doi.org/10.14722/eurousec.2016.23010

personalized services and offers.

Extensive research has been conducted on the perceived benefits and risks of location-based applications, however, the effect of users' understanding of what happens with their location data on privacy behaviour – on usage of LBA and on protection behaviour – is still lacking. This study aims at extending the knowledge about perceived benefits, risks, and knowledge, with data drawn from users' actual experiences. We also propose a novel taxonomy for the risk-benefit calculation that users engage in when using location-based applications. To achieve this, we conducted a set of semi-structured interviews assessing users' beliefs, and connected their knowledge with stated benefits and concerns. This explorative study suggests that there are a number of misunderstandings regarding location-based applications' data use. We find that the participants with limited knowledge of how location-based applications work – or their data privacy aspects – thought more often than others that there are no risks involved in these applications. Better knowledge, on the other hand, seems to be associated with taking more measures to protect oneself from privacy risks. We also find that users who see the most benefits in LBA were the ones who also stated feelings of trust towards different entities, including companies and governmental organizations. On the other hand, the users who stated comments reflecting mistrust mentioned the most concerns over using of LBA. Among the users, surveillance was mentioned most often as a likely risk.

II. Related work

A. Benefits

The location-based applications offer a wide range of benefits to the users. The biggest benefits of these services that were mentioned in a study by Tsai et al. [6] were security or safety related: finding people in an emergency, or tracking the children in one's family, as well as finding information based on one' location. Tang et al. suggest that most location-sharing is purpose driven rather than social driven, such as arranging meetings or transportation [7]. A variety of social applications have gained huge popularity, including services to find a partner nearby, or informing others about one's whereabouts. Sharing information also helps in promoting oneself and enhancing one's status in social circles. How willing one is to disclose location in various situations is influenced by who the requester of information is [8], [9]. Not only closeness to the receiver of the location information, but also trust in the receiving entity decreases privacy concern [10]. Furthermore, trusting beliefs might, in addition to mitigating concerns of privacy risks, increase the users' willingness to disclose information through location-based services [11].

B. Concerns

Users have been found to have particular worries when using mobile devices, and mistrust towards smartphone applications creates agitation in users [12]. These worries include physical damage, data loss, battery life, and lack of trust [12]. In another study, the most likely risks the users see in location-based applications were found to be revealing one's home location, and being stalked (cf. [6]). Also too well targeted advertising seems to create privacy concern and decrease disclosure [13]. Advertising can be seen either as disconcerting or beneficial, depending on the control the user has [14]. The complexity of the topic can be seen in that privacy concern can vary drastically based on the physical situation, or social and technological context [15].

C. Protecting Privacy

Users have several tools at hand to enhance their privacy when using location-based applications. These include switching the location services off altogether, avoiding the usage of services and installation of applications that require one's location, giving access to location information only to certain people and blacklisting others, or location obfuscation, which refers to giving the user an option to share their location at an accuracy that corresponds to their privacy preferences and use case. Users with higher privacy concern take advantage of this functionality and share with lesser precision [16]. In another study, Consolvo et al. found that users tend to share their private information at an accuracy that is most useful to the user [8]. This functionality is not readily available in most systems to date.

In a study by Toch et al. [17] users were found to evolve more sophisticated privacy preferences over time. In another study, users of location-based applications were also found to have difficulties in expressing their privacy preferences [18].

Privacy breaches may have the consequence for an individual to tighten up their privacy protection mechanisms. This was found in a study with undergraduate Facebook users, where privacy violations led to the users having friends-only profiles [19]. Transparent data privacy practices have also shown to decrease users' privacy concern with respect to surveillance [20]. Not only the data privacy practices of companies behind location-based applications, but also governmental legislations have shown to increase feeling of self-control and decrease concern.

D. Misunderstandings

A survey from 2003 by Turow et al. revealed that a vast majority of internet users have overly optimistic views of what happens with their data, and at best, a very limited understanding of data privacy practices [3]. Also the users of location-based applications are often unaware of the data that is collected through the apps they use, and informing them prompts to reevaluate some permissions, or even restrict them [21]. The findings by Turow et al. were repeated by Hoofnagle and Urban in 2012 in a study in which participants' knowledge was tested via a quiz about online advertising [22]. The researchers report that users with high privacy concern seem to have a better understanding of information privacy practices than others. In both these studies, the alarming finding is that the users have an unfounded belief that laws and regulations protect their data from being passed on to third parties. Balebako et al. assessed the gap between users' understandings and actual data leakages and found that users

would like to have more information about data sharing than currently available [23]. Many misunderstandings were revealed within the study, including that the users drastically underestimate how much their data is used for different purposes. There is a gap in the literature in to what extent the limited knowledge affects privacy behaviour in the context of location-based applications.

III. Research Method

To study smartphone users' views and experiences with location-based applications, we conducted a set of interviews. We had an explorative approach, within which we aimed at learning new insights about their experiences and beliefs. The topics covered in the interviews included:

1. Which location-based applications do the participants have? The possibility of some other applications using the location without the users' knowledge was also discussed.
2. Why are the named applications used? What are the benefits the participants see in using location-based applications, and in particular, what kind of benefits have the participants already experienced?
3. What are the reasons for not using some applications?
4. Are the participants aware of any possible risks there might be involved in using location-based applications? Which ones? How did the participants learn about the risks?
5. Has the possible perception of existing risks affected the usage of location-based applications in some way? How exactly?
6. What do the participants believe is done with the users' location information? What do they believe is possible to do with the data? Finally, who is responsible for protecting the user from the possible risks was also discussed with some participants.

Additionally, relationships between the concepts are assessed; we deduct variables from the qualitative interview data to evaluate the relationship of knowledge and privacy behaviour.

A. Data Collection

In total 41 semi-structured interviews were conducted during December 2015 (see Appendix A for the basic interview protocol). This method was chosen because of the explorative nature of the goal of the study, and was expected to yield new insights into the users' awareness of location privacy. Most interviews were carried out in a relaxed atmosphere at the participant's home or in a café when the participants were physically available; otherwise they were conducted through a video call. The interviews were conducted in the participants' native languages, with an exception where the participant was fluent in English. All interviews were audio recorded to obtain verbatim statements from all participants; the participants were asked for consent for this prior to the interview. The transcripts were translated into English prior to analysis by the interviewers, who were fluent both in English and the target language.

B. Participants

The participants were voluntary and recruited from the researchers' extended social circles, while aiming at a good demographic distribution. The requirement for participation was smartphone ownership. Of the 41 recruited participants, 14 were female. The age distribution was slightly skewed towards young adults (M = 29.6 years, SD = 8.8), which is acceptable considering that among this age group, the users can be considered "smartphone dependent", and the smartphone ownership is highest [24]. Thirty-six percent of the participants were students, and 22% worked in the IT sector. The participants represented 14 different nationalities from five continents; the countries represented in the study were Cameroon (1), China (3), Ecuador (1), Germany (19), Hungary (1), Iran (2), Korea (2), Netherlands (2), Peru (1), Spain (2), Sweden (2), Taiwan (3), UK (1), and USA (1). The participants lived in the mentioned countries, and in the cases where the participants were not physically available for a face-to-face interview, these were conducted via video calls.

IV. Qualitative Findings

In analyzing the interviews, we used a mixture of inductive and deductive approaches. As a basis for the codebook, we used existing literature, in particular, the expected benefits and risks found in a study by Tsai et al. [6]. Two independent reviewers coded the interviews, with freedom to be open for new codes during the process.

After the first round of coding, the labels were gathered and grouped into meaningful entities. This round yielded to a revised codebook, which was used by the two independent reviewers for a second round of coding. Finally, the remaining disagreements in the labels were resolved and mutual agreements were reached for each case.

In the following section, we explore the qualitative findings from the interviews. We discuss in detail some of the most important emerged topics, together with some examples. Some of the quotations are translations; we strived to stay as true to the original attitude and choice of words as possible.

A. Applications

We asked the participants what kind of location-based applications they used on their smartphones. We did not check whether the responses were accurate information but concentrated on the participants' views. Some of the participants, however, checked during the interview which applications they have on their smartphones that use their location. At this point, several participants were surprised about some applications using the device's location without their knowledge, however, in each of the cases plausible explanations were found for why the application in question would need the information.

We found navigation to be the most commonly used application type, with almost all participants using it (90%). This category includes maps, navigation aids, as well as apps

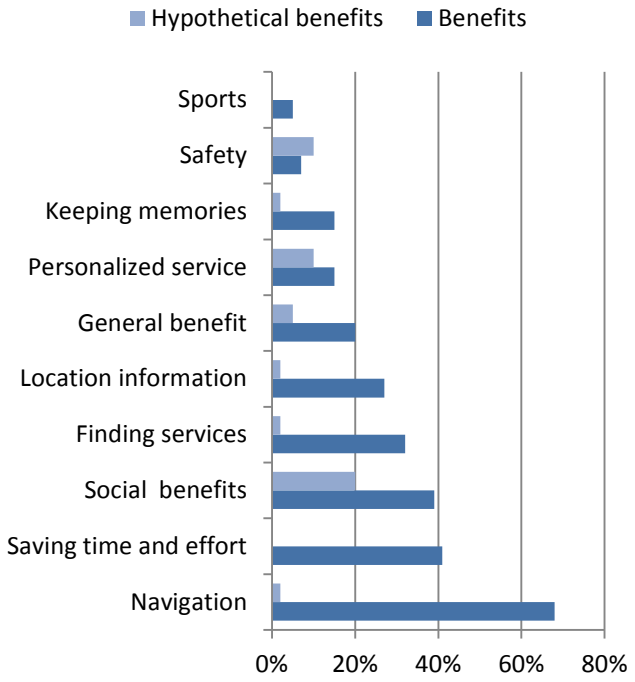

Fig. 1: Benefits and hypothetical benefits received from location-based applications, ordered based on how frequently each category was mentioned.

used specifically for public transportation routes and timings.

The second most mentioned app type was social (54%). This included applications where location services were used for social interactions, such as Facebook, WhatsApp, Twitter, Snapchat, Tinder, Instagram, and Skype. Only applications with location functionalities known to the user were considered. However, the location functionalities were not used in all cases.

Twenty-two percent of the participants had applications that they used for finding services, including Yelp, Booking.com, Airbnb, and others. Fifteen percent mentioned using a weather app with location functionality.

The other applications mentioned to be in use were different applications for sport activities (10%), safety applications such as a "find my phone" app (5%), as well as taxi and ride sharing applications (5%). Finally, other location-based applications not included in the above mentioned categories were mentioned by ten participants (24%). These include music streaming, fashion and shopping applications that the participants stated use their location.

B. Benefits

We asked the participants about the benefits they have experienced with location-based applications. The participants also mentioned benefits that they could imagine existing, or benefits that they believe their friends or family have experienced. We labelled the comments of this latter type as hypothetical benefits to make the distinction between actual benefits and the ones that the participants have not experienced themselves (cf. Fig. 1).

The most commonly stated benefit was, perhaps unsurprisingly so, *navigation* (71%). Forty-two percent of the participants said that location-based applications have helped them in *saving time and effort*, mostly by simplifying the interaction by requiring less user input. *Social benefits* were also mentioned by 42% of the participants. These included sharing one's location in a group when setting up meetings, for safety reasons for example in the case of elderly family members or ones with memory issues, location-based gaming, or for social recognition. Social recognition was also mentioned several times, though only as a hypothetical benefit, as stated by participant 18 as follows: *"Well, I think this kind of location-sharing app is commonly used by those who want to show off. Those people can share wherever they are [visiting] for example, some fancy, high-class restaurant or going somewhere few people are able to go."* Social benefits were mentioned as a hypothetical benefit by altogether eight participants (20%). The reason for that social recognition was seen only as a hypothetical benefit could be that it might not be socially acceptable to be showing off, and as a consequence, it is safer to avoid talking about it in active voice.

One third of the participants mentioned a benefit of *finding services*, such as stores, restaurants, or accommodation (32%). Other *location Information*, including store opening hours, or information regarding a currently visited point of interest, were mentioned by roughly quarter of the participants (27%). *Personalized service* was mentioned as a benefit by 15 per cent. These included search results that fit to the users' context, or adverts and promotions based on their location. Four participants thought this could hypothetically be beneficial (10%).

Six participants mentioned *keeping memories* as a benefit (15%). In these cases, location traces would be used mostly as something like diary entries.

A few participants found *safety* features a benefit from LBA (7%). Mentioned benefits were about finding one's family members or stolen property. Safety was mentioned also as a hypothetical benefit (10%), mainly for being able to track family members who need to be taken care of (such as kids or elderly). Also the possibility for the government to track citizens for safety reasons was brought up.

There were two mentions of *sports* as a benefit, including running and biking (5%). Finally, *general benefit* was mentioned by eight participants (20%) including benefits such as convenience, making one's life better, "connecting the physical world with the virtual world", or providing a benefit for the society by creating more data.

Also, one tenth of the participants (10%) expresses that the data will be used somehow to develop or improve the services. "*For most developers, collecting user information is very important to help improve the quality of service*" (P7). This is however not clearly a benefit from using the applications.

C. Risks and Hypothetical Risks

The participants were asked whether they thought there were some risks involved with using location-based applications. The participants talked mostly about actual risks that one should be aware of, but quite often also hypothetical

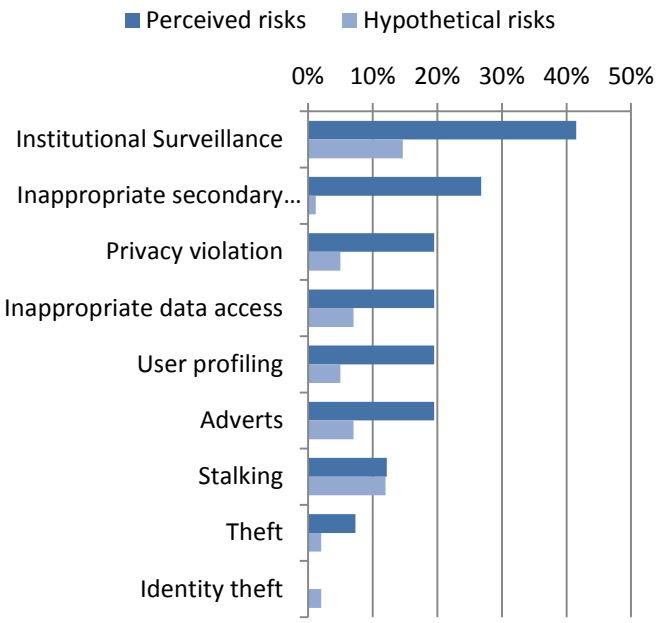

Fig. 2: Risks and hypothetical risks by the number of participants having mentioned them. A risk is called hypothetical if the participant was not currently concerned about it but mentioned it as a hypothetical scenario.

risks were mentioned. These are risks that are considered possible in some circumstances but are not seen at all likely to happen to oneself, at least with the current status quo. We differentiate between these two by dividing them into separate categories: risks, and hypothetical risks. For an overview, see Fig. 2.

The risk mentioned most often was institutional surveillance – mentioned by 42% of participants. This category includes statements regarding the police or the state following one's actions. Opinions such as the following were stated: "*I'm afraid of the state, institutions, police, that they draw conclusions and predict 'Minority Report' style things, and classify you. That's disgusting, that's the problem because I think the private sector is not the problem.*" (P3), as well as simply: "*[…] it's very easy to spy on people.*" (P30) or "*There is no better way to control people*" (P21). Surveillance was also mentioned by six participants (14.6%) as a hypothetical risk, which means that these people talk about surveillance as a possible risk but do not feel threatened by it, and, more often than not, the possibility leads to no action with respect to privacy behaviour. "*If I was on the run and someone with access to the data wanted to find me, then yes, it would be possible to find me. But I'm not on the run since I don't have any bad things going on*" (P27).

The second most reported concern was about *inappropriate secondary use of data* (27%). The participants expressed concern over not knowing what is done with their data, or that it even might be used by people with bad intentions, or sold to third parties without one's knowledge. The hypothetical risk of inappropriate secondary use of data was expressed by the same concerns, with the difference of not feeling directly threatened by them. "*Maybe at this point in time not yet, but in the future it might be quite dangerous. I mean you never know if someone has a good or bad intention with this data and who they might sell this data to*" (P15). Slightly fewer participants reported concern of *inappropriate data access* (20%). While inappropriate data access refers to a situation where the user's information has gotten in the hands of parties not originally intended, the inappropriate secondary use of data specifies that the information is also used for purposes not originally intended by the data subject, nor permitted by them. A typical comment reflecting inappropriate data access would be, as stated by P15: "*Well, the fact that these companies know where you are located and maybe they, I don't know how it works, but maybe some hackers or someone that's good with programming can actually also get this information.*" Three further participants talked about a hypothetical risk of inappropriate data access.

Adverts were worrying eight of the participants (20%), and a connected user profiling also by eight. In some cases the participants combined the concepts; however, mostly what was mentioned was either about adverts or about user profiling. The concepts are closely connected, as adverts here refer to behavioural advertising, which is done based on the user profiles. Altogether either adverts or user profiling was mentioned by 24%.

Privacy violation was brought up also by eight participants, possibly in lack of a more precisely directed concern. The statements included comments such as the following: "[…] *if talking about the risk, I think it's about the user's privacy. For example, I think it's personal information, it's private information*" (P19).

Stalking was mentioned by five participants as a risk. "[…] *you are really transparent, which makes stalking much easier*" (P41). In some cases, concern of stalking was seen in hacking, for example, "*people could stalk you if they hack the app*" (P11). Further five participants saw stalking as a hypothetical risk. Even if the participants are not concerned, they are still aware of the possibility, for example, P27 mentioned the concern as follows: "[…] *if my boss could see where I am... Then he could have seen that I'm at a job interview somewhere else. But that takes that there is another person spying on you, otherwise is not a risk. So I don't really think there are any risks.* "

The concern of theft results from the possibility of getting tracked, through self-reports of where you are, for example via social media. "[…] *it is a well-known risk that it is not always good to tell everyone where you are all the time. It is not always that good, thefts for example*" (P28). Some were also concerned about disclosing home location through tracking: "[…] *so he can track where my home is, he can steal my posts*" (P1). This category covers two different types of theft: firstly, the risk of thieves getting to know where there is an empty apartment (for example, because of holiday posts on social media), and secondly, of robbery after an individual's whereabouts have been figured out through the use of location-based applications. Once also a hypothetical risk of theft was mentioned, and it also falls under these categories. Identity theft was mentioned only as a hypothetical risk; it did

not come up as a potential risk that someone would be currently concerned about.

D. Trusting Beliefs

Tradeoff. Half of the participants felt that there is a tradeoff in using location-based applications (49%). "*I take that risk, because I get something instead. But that is the limitation. I need to get something in return, so that I divulge my location*." (P24).

There is no risk. Almost half of the participants were of the opinion that there is no risk involved (44%). The vast majority of these stated that only dishonest people have some risks and notably, that "they have nothing to hide": "*Personally I don't care much. Got nothing to hide. The benefits are more than the inconvenience.*" (P8). Others thought that when too much data is being collected, it cannot be used anymore for anything useful – thus there are no risks in data collection.

Powerlessness. Many reported powerlessness over their data (42%). These participants are not comfortable about how their data is being handled, or about the lack of control thereof. As an example, P31 commented on the topic of reading app permissions as follows: "[…] *the data is stored, but there are so many updates and partially you have to accept various conditions with it, and I don´t know anyone who is reading them carefully and dealing with them. So I think that one quickly loses track of all these functions and updates that you installed on a daily basis, without looking what is now really changed since, I think that is not transparent.*"

Trust. Trusting comments were stated by every third participant (29%). In many of these comments the participants stated that they trust that companies, in particular big companies such as Google and Facebook, treat their data correctly. "*Google is a world-known company, which means they have the obligation to protect the customers' data. Sometimes those apps asking users' location only use it for their servers, so no need to worry about that.*" (P17). P28 commented on Google: "*I don't think that they would sell the information. That would be bad PR for them.*" Several comments also reflected trust in the governmental organizations' data privacy practices.

Mistrust. Comments were labelled as mistrusting when they reflected that the participant did not believe the companies or government organizations are honest about data privacy practices, or when there were feelings that data is being unnecessarily saved or used. These were slightly less common than those stating trust (24%). An example of a statement showing mistrust would be the following by P24: "*Well if you hear how all the big companies like Facebook and Google pass information to security agencies… Then I think that they are not able to protect my own privacy.*"

E. Knowledge

Various comments within the interviews included statements that reflect either misunderstandings of different types with respect to LBA, or a good knowledge of the data flow of LBA or technical understanding of how LBA work.

Table 1: Protective measures taken against privacy risks on LBA. The categories are partly overlapping, meaning that some participants use more than one protective measure.

Protective Measures	Percentage
Technical measures	53.7%
Avoiding usage	39.0%
Educating oneself	24.3%
No measures taken	20.0%

The categories are partly overlapping as some participants shared comments showing good knowledge, and on other statements, some misunderstandings. Knowledge was not systematically recorded for all the participants but rather, the issues came up during the interviews. Studying the extent of knowledge of information flows in LBA systematically remains thus a topic for future studies.

Limited knowledge. During the interviews, we found a majority of the participants having misconceptions about LBA. Altogether 25 participants (61%) had some limitations in their knowledge. These could be further divided into subcategories:

1. Misunderstanding about some technical detail. The most frequently recorded misunderstanding was that GPS would be the only way of finding out one's location, and by switching GPS off, the phone's location could no longer be tracked.
2. Statements where a participant says that they are not fully aware of how things work.
3. Misunderstandings regarding what would be done with the data. For example, some participants were convinced that user profiles are not being used by third parties, or that information is not used because that would be too much effort: "*I think they won't spend so much effort in combining the data?"* (P13).

Good knowledge. In this category, we included comments that showed good knowledge of how LBA work, for example with respect to what is possible to find out based on the data, or of data protection regulations. "*I know that there are certain laws that state how long such information can be saved*" (P36).

F. Protective Measures

The participants explained what kind of measures they take when they are somehow concerned or see some risks in using location-based applications. The comments regarding protective measures are divided into four categories as follows (c.f. Table 1). The categories are partially overlapping because some participants mentioned more than one such reason.

Technical measures. The most popular protective measures category, technical measures, combines all technical possibilities that were mentioned being used to protect oneself, such as switching off location services, or denying location access for some applications. "I turn my location off when I don't need it" (P9). Others mentioned defining their location settings as a privacy-protection method: "I tick of who is allowed to use it and who isn´t" (P25). The above quotations are typical statements we recorded for protecting measures in

a technical context, mentioned by 54% of the participants.

Avoiding usage. 39% of the participants reported avoiding usage, with varying degree of clear privacy reasons. Since we labelled into this category also comments that did not explicitly mention concern, this cannot be taken purely as a measure to protect oneself from privacy concern. A number of privacy-related statements were recorded, including: "*[…] when I got the impression that an app which is not at all related to "location", but is asking for it, then it is enough of a reason for me not to download that app.*" (P24). Twenty percent of the participants stated explicitly that they avoid using LBA due to privacy reasons. Other reasons for not using LBA, or avoiding their usage, included not seeing benefits in these applications (37%), annoyance (5%), and technical reasons such as saving battery (10%). These reasons either referred to a single application, or to location-based applications in general. Insome cases, participants even stated that they do not have privacy concerns.

Educating oneself. Roughly a quarter of the participants expressed statements we covered with the category educating oneself. Exemplary here are comments that one reads the terms of agreement or checks for certain permissions before downloading an application.

No measures taken. Finally, one fifth of the participants remarked that they do not take any measures to protect themselves. As an example, participant P3 discussed about data protection and companies knowing where he has been through geotagged pictures as follows: "*I feel uncomfortable. I know that, but I kind of ignore it. It is somehow worth it.*"

G. Source of Information and Responsibility

Some participants reported where they had learned about the data use and risks involved in location-based applications. According to these participants, media was the main source of information, including television, newspaper, radio, and online articles. The information about risks and data misuse comes mostly through reports of scandals. Other mentioned information sources, though playing only a small role, were through work, friends, being self-learned, and other sources.

Some comments were given as to whose responsibility it is to protect the users from privacy breaches. Such comments were recorded only from 27% of the participants. Some users saw that the user is responsible for the data protection. A typical comment stating users' responsibility was that by P31: "*I am responsible for what data I would like to give away, so I can also switch off all the location services and only the network provider knows in which area my phone logs in.*"

Even more frequently participants were of the opinion that the state would need to take the responsibility of protecting users. This was stated almost unanimously among the participants who took a stand on whose responsibility data protection is. P32 said: "[…] *I don´t really believe that every user has the overview and would be able to protect oneself adequately. I don´t think the App Store as a resell and download platform is the right contact person. Neither are the network providers, because they have nothing to do with the apps. In my opinion the government is responsible for regulating with laws or at least some rules for the app providers what they are allowed to do and what are not.*"

H. Other Variables

Here, we present how we deduced variables from the interview data to conduct further analysis.

1) Knowledge

Statements that reflected either good knowledge regarding the functionality of location-based services – or the lack thereof – were partially overlapping. This means that a participant showed good understanding with some comment, and limitations of knowledge could be seen in some other comment by the same participant. We took all comments reflecting either end of this spectrum, and created a new variable called knowledge. This variable measures knowledge on a five-point scale:

1. Limited knowledge (the participant has at least one comment showing limited knowledge, but none showing good knowledge).
2. Both kinds of comments are present, but there are more stating limited than good knowledge.
3. There are as many comments stating good knowledge as limited knowledge, at least one each.
4. Both kinds of comments are present, but more reflect good knowledge than limited knowledge
5. Good knowledge (at least one comment reflecting good knowledge and none of limited knowledge).

On this new scale, mean was 2.85, and standard deviation 1.67. Seventeen percent of the participants were not categorized because of lack of comments that could be used to categorize them, thus, they are excluded from the analysis related with knowledge. Some of the limitations are more severe than others and thus have unequally big consequences on privacy behaviour; the same applies also for good knowledge. Taking these differences into account is out of the scope of this work and a topic for future research.

2) Benefits and Risks

We created a variable listing the number of different types of benefits that were seen in using LBA to quantify the perceived usefulness of LBA. The median number of benefits mentioned was two (M = 2.73, SD = 1.57). Similarly, we counted the amount of different risks that are seen in using LBA and introduced a variable that lists the sum. Also for this variable the median was two (M = 1.63, SD =1.55).

We created six binary variables of whether or not the most commonly mentioned risks were mentioned by the participant. We considered only the actual risks, and not the hypothetically mentioned ones. The considered risks were *surveillance, secondary use of data, privacy violation, inappropriate data access, user profiling,* and *adverts.*

3) Trust

We created a variable to measure trust similarly as to measure knowledge (cf. Section IV.H.1). On this five-point scale ('1' representing most mistrusting, and '5' representing most trusting) the mean trust score was 3.21 (SD = 1.90). Altogether 46.3% of the participants (19 individuals) were

given a trust score; others did not state comments that could be regarded either as trusting or mistrusting. While it could be argued whether trust can be considered a trait, we consider the participants who stated mostly trusting comments as *trusting individuals*, whereas the participants whose comments reflected mostly mistrust, we call *mistrusting individuals.*

V. Results

In an attempt to find out about the relationships between the various concepts found in the data, we ran tests using the statistical tool SPSS. Our goal was to gain some insight to how knowledge and beliefs affect users' privacy behaviour. As our variables were not systematically measured from all participants, these results should be taken rather as directive, than conclusive. While we cannot state anything about causality, we did find some relationships between these concepts.

A. Knowledge

We looked at the association between knowledge and taking protective measures against privacy risks. A Mann-Whitney U-test showed that the participants who stated avoiding usage of location-based applications have a significantly higher knowledge score than those who do not ($U = 84.0$, $p = .043$). Furthermore, statements implying that there are no risks involved were stated significantly more frequently by participants with lower knowledge scores ($U = 66.5$, $p = .006$).

We also found that the users who felt that there was no risk to their privacy associated with using location-based applications did not take technical measures to protect themselves, $X^2(1, N = 41) = 5.33$, $p = .023$.

B. Trust and Mistrust

A nonparametric correlation test showed a moderate positive correlation between trust and the number of benefits seen (r_s (17) = .449, $p = .027$). On the other hand, a moderate negative correlation was found between mistrust and the number of risks seen (r_s (17) = -.395, $p = .0479$)

The users who take technical measures to protect their privacy when using location-based applications are significantly more mistrusting than those who do not, $U = 18.00$, $p = .027$. Similar effects were not found with educating oneself, nor with avoiding usage.

Some participants stated that it is a tradeoff to use location-based services – mostly a tradeoff between receiving benefits and losing privacy. The participants who talked about a tradeoff were more concerned about surveillance than others, $X^2(1, N = 41) = 4.19$, $p = .042$.

VI. Taxonomy

Assuming that a user of location-based services engages in a cost-benefit calculation to define whether or not the benefits of using a given service outweigh the possible risks, we propose a first step towards a taxonomy of cost-benefit calculation in the usage of location-based applications (cf. Fig. 3). The calculation consists of perceived risks and perceived benefits. Different categories are expected to have different weighs in the calculation. The categories were validated in a small user study with participants who were not familiar with this study ($N = 8$).

Within the perceived benefits, five categories were identified: saving time and effort, social benefits, safety, finding information, personalized services, and quantified self. Monetary benefit was not mentioned in the interviews, and its inclusion in the taxonomy is a topic for future research.

The perceived risks can be further divided into related categories – three such categories were identified. These include surveillance, privacy violation, profiling, and criminal activities. Privacy violation includes surveillance, inappropriate data access and inappropriate secondary use of data, as well as other cases where the user feels that their privacy has been violated. Inappropriate data access is an inevitable first step before inappropriate secondary use of data, however is not necessarily followed by it. Profiling consists of user profiling, and the related behavioural advertising. Behavioural advertising was identified also as a benefit based on the interview data, and is the only item that is found on both the sides of the calculation. Finally, criminal activities include stalking, theft, and identity theft.

VII. Discussion

In this work we conducted qualitative interviews on the usage of location-based applications (LBA), and propose a taxonomy based on the findings. The taxonomy considers the found perceived benefits and perceived risks as input parameters for a cost-benefit calculation when users make a privacy decision of whether or not to engage in the usage of such an application. The other findings from the study are discussed here.

A. Misunderstandings

There were several misunderstandings found about how LBA work. The most common misunderstanding was that by switching the GPS off, one's location could not be tracked anymore. The participants with less knowledge also turned out not to protect themselves from privacy risks by avoiding usage of LBA as much as those who did not have such limitations in their knowledge. It seems plausible that users with such limitations did not think there is privacy risks involved in using LBA, and as a consequence, did not see any reason to avoid using them. The connection could be also seen in that users, who said that there are no risks involved, also did not use technical measures, such as switching the location services off or uninstalling applications, as often as others. Location blurring was not mentioned by any participants as a protection method – perhaps the option is still not that readily available. Some individuals stated privacy concern and also admitted that their knowledge is still limited. Nowadays, having some basic understanding of the information flow, or at least of the possible risks, is a precondition to being in control of one's personal information. The results of this study suggest that the user should have less responsibility and be adequately protected even without extensive knowledge about data privacy issues.

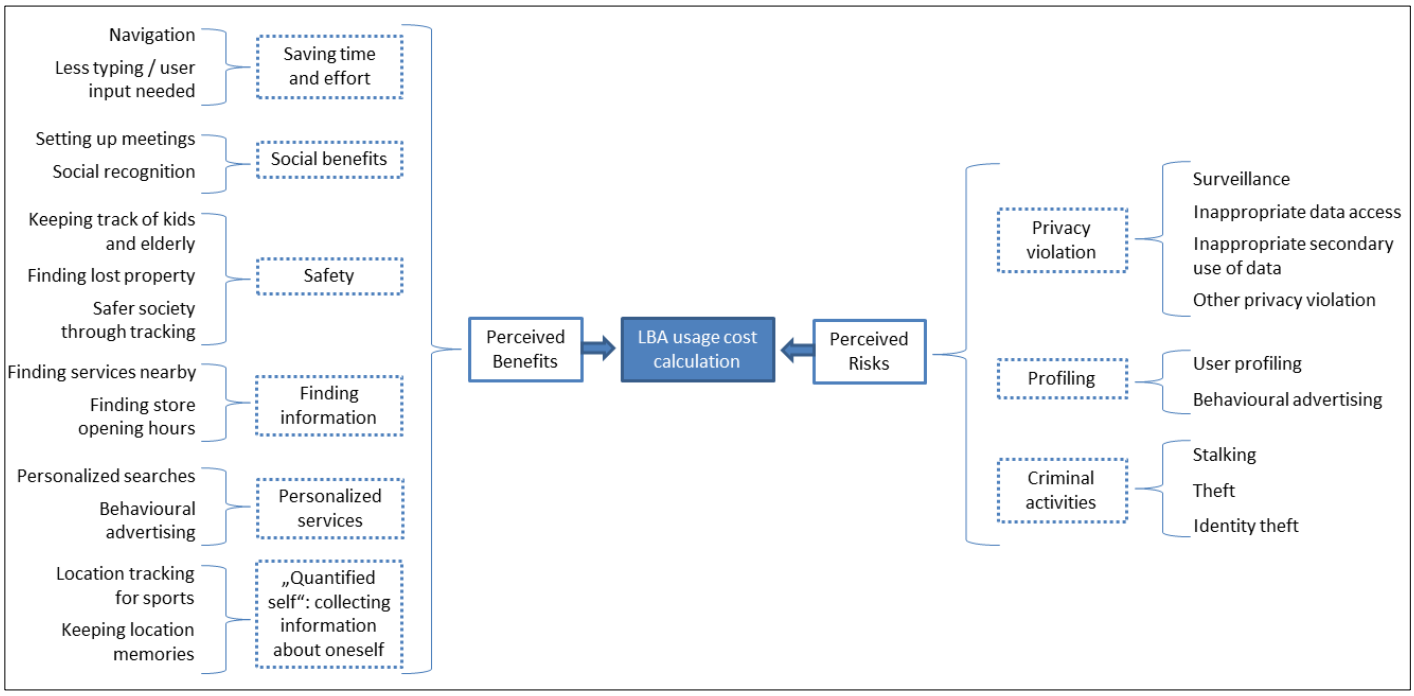

Fig. 3: Proposed taxonomy for cost-benefit calculation in the context of usage of location-based applications. The calculation is done based on perceived benefits and perceived risks. The first can be divided into five, and the latter into four categories.

A majority of the participants who said that there is no risk also said that they "have nothing to hide". This statement has been discussed in recent literature: Solove discussed the concept stating that often the users who say they have nothing to hide have a very myopic view of what "privacy" means, understanding it merely as secrecy – hiding something bad [25]. Also our study supports the assumption that the "nothing to hide" view could be a consequence of a myopic view and limited knowledge. Our results show that there is a difference between the knowledge scores of the users who state that there are no risks involved in using LBA: the users who think there are no risks have a significantly lower knowledge score than those that do not. We also find that the users who avoid using location-based applications, for privacy reasons or otherwise, have higher knowledge scores than others. It could be that the users with better knowledge are more aware that there might be some risks involved, and as a consequence they avoid using location-based apps, or use technical measures. This is, however, a speculation and cannot be directly inferred from the data. The interpretation is nevertheless in line with an earlier finding that the internet users who can be categorized as privacy fundamentalists based on the Westin categorization [26] also have a better understanding of what happens with the data [22]. However, it has been suggested that other instruments might provide better options than this categorization [27].

B. Perceived Risks

In an earlier study, the most salient risks in using LBA were reported to be revealing one's home location, and getting stalked [6]. These particular concerns came up also in our study, but these were some of the least mentioned ones. The most frequently mentioned risks in this study were surveillance and secondary use of data. Also rather often mentioned issues were a general privacy violation, inappropriate data access, user profiling, and adverts.

We would also like to point out two distinctive cases of perceived risks – the risk of location information being inappropriately accessed or used by individuals, or by companies and institutions. The concerns categorized as surveillance or profiling include a worry of the data being accessed by organizations, whereas statements categorized as privacy violation and criminal activities reflect worries that the information is finally used by unauthorized individuals.

C. What Is Done With Location Information?

What do users think happens with the data when they use LBA? Majority of the participants stated that it is used for user profiling, and half mentioned that it is sold to third parties. This is not to say that the rest of the participants did not think that profiles are created or data is sold – they just did not mention it within the interviews. These results also do not take a stand on whether the participants thought the practices are beneficial or harmful.

D. Protective Measures and Avoiding Usage

The most important reason for not using LBA was stated as not seeing benefits in the usage. Privacy concern seemed to also be an important reason for many; approximately one fifth of the participants stated privacy issues as the reason for not using LBA. It seems that more often than avoiding usage to protect themselves from privacy risks, the users take some technical measures. This includes turning the location service off, or even uninstalling applications. This was particularly typical for users showing mistrust towards different organizations, including governmental organizations and companies. Privacy measures such as blacklisting people, or

location obfuscation, were not mentioned by our interview participants. It can be that these options are not readily available in most applications that the participants use. Avoiding usage of particular applications, or location-based applications altogether, was mentioned by nearly 40% of the participants; however, not all of these are necessarily for purely privacy reasons.

While an important reason for not using location-based applications is not getting benefits out of the usage, many of those who still continue using the LBA find that there is a tradeoff, and one has to compromise privacy to get a benefit. The feelings of tradeoff were in particular associated with concerns of surveillance. Often also powerlessness over one's data was expressed. Both these statements suggest that there is not enough transparency, and users do not know whom to trust.

E. Who Protects?

Whose responsibility is it finally to care for end-user privacy? This topic has been previously discussed by Cottrill with a review of legal, technological, and practical aspects of protection [28]. In our study the topic was not discussed by all the participants, however, nearly all of these stated that it is indeed the state's responsibility. In this study we heard also several comments of mistrust towards data protection laws, and in particular, towards the potential big brother effects that could ensue. In earlier studies, government regulations have shown to increase trust [29] – but the condition for this might be an adequate base level of trust towards the governmental data privacy practices.

VIII. Conclusions

Our most important results are qualitative findings from 41 interviews conducted with participants from various countries. Our results suggest that a large number of users of location-based applications have overly optimistic views about what is done with the users' data, and that the limitations on knowledge are often associated with statements that no risks are included in using location-based applications. The lacking risk perception could be an explanation to why users with limited knowledge were also found to take fewer measures to protect themselves from privacy risks when using these applications. We also find a sizable user segment that is mistrusting towards companies and governmental organizations, which is associated with seeing more risks in using location-based applications and with using protection mechanisms against privacy risks. We also identified a prominent feeling of a tradeoff accompanied with using location-based applications – the users think there are risks, but accept them as a price they have to pay when using these services. Our findings suggest that in particular, for the user segment with limited knowledge, an adequate level of privacy protection should be provided also without explicit user action.

IX. References

[1] L. Palen and P. Dourish, "Unpacking 'privacy' for a networked world," *Proc. Conf. Hum. factors Comput. Syst. - CHI '03*, no. 5, p. 129, 2003.

[2] S. Petronio and W. T. Durham, "Communication privacy management theory.," *Engaging theories in interpersonal communication: Multiple perspectives.* 2008.

[3] J. Turow, L. Feldman, and K. Meltzer, "Open to Exploitation: America's Shoppers Online and Offline," *Annenb. Public Policy Cent.*, p. 10, 2005.

[4] A. McDonald and L. F. Cranor, "The Cost of Reading Privacy Policies," *I/S - A J. Law Policy Inf. Soc.*, vol. 4, no. 3, pp. 1–22, 2008.

[5] C. Jensen and C. Potts, "Privacy policies as decision-making tools," *Proc. 2004 Conf. Hum. factors Comput. Syst. - CHI '04*, vol. 6, no. 1, pp. 471–478, 2004.

[6] J. Y. Tsai, P. G. Kelley, L. F. Cranor, and N. Sadeh, "Location-Sharing Technologies : Privacy Risks and Controls," *A J. Law Policy Inf. Soc.*, vol. 6, no. 2, pp. 119–151, 2010.

[7] K. Tang, J. Lin, and J. Hong, "Rethinking location sharing: exploring the implications of social-driven vs. purpose-driven location sharing," *Proc. 12th ACM Int. Conf. Ubiquitous Comput. - Ubicomp '10*, vol. 12, no. 4–5, pp. 85–94, 2010.

[8] S. Consolvo, I. E. Smith, T. Matthews, A. LaMarca, J. Tabert, and P. Powledge, "Location disclosure to social relations," in *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '05*, 2005, p. 81.

[9] J. Venkatanathan, J. Lin, M. Benisch, D. Ferreira, E. Karapanos, V. Kostakos, N. Sadeh, and E. Toch, "Who, when, where: Obfuscation preferences in location-sharing applications," *ISR Technical Reports 2011, CMU-ISR-11-110, Carnegie Mellon University*. pp. 1–12, 2011.

[10] H. J. Smith, T. Dinev, and H. Xu, "Theory and Review Information Privacy Research: an Interdisciplinary Review 1," *MIS QuarterlyInformation Priv. Res.*, vol. 35, no. 4, pp. 989–1015, 2011.

[11] H. Xu, H. Teo, and B. C. Y. Tan, "Predicting the adoption of location-based services: the role of trust and perceived privacy risk," *Proc. 26th Int. Conf. Inf. Syst. (ICIS 2005), Las Vegas*, no. Beinat 2001, pp. 897–910, 2005.

[12] E. Chin, A. P. Felt, V. Sekar, and D. Wagner, "Measuring user confidence in smartphone security and privacy," *Proc. Eighth Symp. Usable Priv. Secur. - SOUPS '12*, no. 1, p. 1, 2012.

[13] A. Goldfarb and C. Tucker, "Online Display Advertising: Targeting and Obtrusiveness," *Mark. Sci.*, vol. 30, no. 3, pp. 413–415, 2011.

[14] M. Benisch, P. G. Kelley, N. Sadeh, and L. F. Cranor, "Capturing location-privacy preferences: Quantifying accuracy and user-burden tradeoffs," *Pers. Ubiquitous Comput.*, vol. 15, no. 7, pp. 679–694, 2011.

[15] L. K. John, A. Acquisti, and G. Loewenstein, "Strangers on a Plane: Context-Dependent Willingness to Divulge Sensitive Information," *J. Consum. Res.*, vol. 37, no. 5, pp. 858–873, 2011.

[16] M. Poikela, R. Schmidt, I. Wechsung, and S. Möller, "Locate!-When do Users Disclose Location?," in *Workshop on Privacy Personas and Segmentation (PPS) at the Tenth Symposium On Usable Privacy and Security (SOUPS)*, 2014.

[17] E. Toch, J. Cranshaw, P. H. Drielsma, J. Y. Tsai, P. G. Kelley, J. Springfield, L. Cranor, J. Hong, and N. Sadeh, "Empirical models of privacy in location sharing," *Proc. 12th ACM Int. Conf. Ubiquitous Comput. - Ubicomp '10*, p. 129, 2010.

[18] N. Sadeh, J. Hong, L. Cranor, I. Fette, P. Kelley, M. Prabaker, and J. Rao, "Understanding and capturing people's privacy policies in a mobile social networking application," in *Personal and Ubiquitous Computing*, 2009, vol. 13, no. 6, pp. 401–412.

[19] F. Stutzman and J. Kramer-Duffield, "Friends only: Examining a privacy-enhancing behavior in Facebook," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2010, pp. 1553–1562.

[20] A. Oulasvirta, T. Suomalainen, J. Hamari, A. Lampinen, and K. Karvonen, "Transparency of intentions decreases privacy concerns in ubiquitous surveillance," *Cyberpsychology, Behav. Soc. Netw.*, vol. 17, no. 10, pp. 633–638, 2014.

[21] H. Almuhimedi, F. Schaub, N. Sadeh, I. Adjerid, A. Acquisti, J. Gluck, L. Cranor, and Y. Agarwal, "Your Location has been Shared 5,398 Times! A Field Study on Mobile App Privacy Nudging," *Proc. 2015 ACM Conf. Hum. factors Comput. Syst.*, pp. 787–796, 2015.

[22] C. J. Hoofnagle and J. M. Urban, "Alan Westin's Privacy Homo

Economicus," *Wake Forest Law Rev.*, vol. 49, no. 2, pp. 261–317, 2014.
[23] R. Balebako, J. Jung, W. Lu, L. F. Cranor, and C. Nguyen, "'Little Brothers Watching You': Raising Awareness of Data Leaks on Smartphones," *SOUPS '13 Proc. Ninth Symp. Usable Priv. Secur.*, pp. 12:1–12:11, 2013.
[24] A. Smith, K. Mcgeeney, L. Rainie, and S. Keeter, "U.S. Smartphone Use in 2015," *Smartphone Differ.*, p. 60, 2015.
[25] D. J. Solove, "'I've Got Nothing to Hide' and Other Misunderstandings of Privacy," *San Diego Law Rev.*, vol. 44, pp. 1–23, 2007.
[26] P. Kumaraguru and L. Cranor, "Privacy indexes: A survey of westin's studies," *Science (80-.).*, vol. Tech. rep., no. December, pp. 1–22, 2005.
[27] S. Preibusch, "Guide to measuring privacy concern: Review of survey and observational instruments," *Int. J. Hum. Comput. Stud.*, vol. 71, no. 12, pp. 1133–1143, 2013.
[28] C. D. Cottrill, "Location privacy: Who protects?," *URISA Journal-Urban Reg. InformationSysmetms Assoc.*, vol. 2, no. 23, p. 49, 2011.
[29] C. W. Thomas, "Maintaining and Restoring Public Trust in Government Agencies and their Employees," *Adm. Soc.*, vol. 30, no. 2, pp. 166–193, 1998.

Appendix

A.1 Interview Script

- How many location-sharing applications do you have on your smartphone?
  - Which ones?
- Which other applications do you have?
  - Are there some applications that potentially use location features without your knowledge?
- Why are the mentioned location-based applications being used?
- If you do not use location-based applications, why not?
- What are some possible benefits you think there are in using location-based applications?
- What kind of benefits have you already had?
- Have you heard of any possible risks that there might be?
  - What risks?
  - How have you heard about the risks?
- How has the knowledge of possible risks affected the use of location-based applications?
  - Have you chosen not to install some applications?
  - Have you used applications less or differently because of the knowledge?
- What do you think is done with your data?
- Do you believe the companies that create location-based applications can access your location data?
- What do you believe the companies do with the location data?
- What do you believe is possible to do with the location data?
- How likely do you believe it might be that…
  - …your home or work address becomes known?
  - …the data is collected to be sold to third parties such as advertisers?
  - …the data is collected to create a profile of you?
  - …the data is combined with other information to create a profile of the user? …and sold to advertisers?