“You just can’t know about everything”: Privacy Perceptions of Smart Home Visitors

Karola Marky
TU Darmstadt
Darmstadt, Germany
marky@tk.tu-darmstadt.de

Sarah Prange
Bundeswehr University
Munich, Germany
sarah.prange@unibw.de

Florian Krell
TU Darmstadt
Darmstadt, Germany
krell@tk.tu-darmstadt.de

Max Mühlhäuser
TU Darmstadt
Darmstadt, Germany
max@tk.tu-darmstadt.de

Florian Alt
Bundeswehr University
Munich, Germany
florian.alt@unibw.de

ABSTRACT

IoT devices can harvest personal information of any person in their surroundings and this includes data from *visitors*. Visitors often cannot protect their privacy in a foreign smart environment. This might be rooted in a poor awareness of privacy violations by IoT devices, a lack of knowledge, or a lack of coping strategies. Thus, visitors are typically unaware of being tracked by IoT devices or lack means to influence which data is collected about them. We interviewed 21 young adults to investigate which knowledge visitors of smart environments need and wish to be able and protect their privacy. We found that visitors consider their relation to the IoT device owner and familiarity with the environment and IoT devices when making decisions about data sharing that affect their privacy. Overall, the visitors of smart environments demonstrated similar privacy preferences like the owners of IoT devices but lacked means to judge consequences of data collection and means to express their privacy preferences. Based on our results, we discuss prerequisites for enabling visitor privacy in smart environments, demonstrate gaps in existing solutions and provide several methods to improve the awareness of smart environment visitors.

CCS CONCEPTS

• **Security and privacy** → *Privacy protections*; • **Human-centered computing** → **Empirical studies in ubiquitous and mobile computing**.

ACM Reference Format:
Karola Marky, Sarah Prange, Florian Krell, Max Mühlhäuser, and Florian Alt. 2020. “You just can’t know about everything”: Privacy Perceptions of Smart Home Visitors. In *19th International Conference on Mobile and Ubiquitous Multimedia (MUM 2020), November 22–25, 2020, Essen, Germany*. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3428361.3428464

1 INTRODUCTION

Ever more people integrate IoT devices into their households [4, 48] and living in a smart home will become the norm. Furthermore, public spaces, such as hotel lobbies, are also increasingly equipped with IoT devices. This also means that people will be more likely to encounter smart environments in the role of a *visitor*, for example, as they visit public or private places in which they are not residing.

IoT devices can harvest information of persons in their surroundings and visitors are no exception. The privacy of visitors is at risk for several reasons: they often lack knowledge about IoT devices, they are not aware how IoT devices might violate their privacy, or they lack means to influence which data is collected [42, 56].

Prior work concluded that visitors require means to exert control over the data collected about them and showed that visitor privacy needs to be explored further [56]. Due to different configuration options of smart environments, visitors today mainly have three “extreme” options: 1) refraining from visiting the smart environment, 2) switching off the IoT device which might result in a tension with the owner and limits access to the device’s features, or 3) sacrificing their privacy. Neither solution is practical since it places an undesirable burden on visitors and IoT device owners.

We contribute an understanding of how aware visitors are of the data being collected by IoT devices in their vicinity, how smart environments affect the behaviour of visitors, and what kind of information they need. We then suggest applicable measures to protect visitors’ privacy, focusing on existing IoT devices. We focus on lack of know-how, showing that such knowledge gaps prevent visitors from protecting their privacy. In particular, we shed light on the following research questions:

- **RQ1:** Which aspects do visitors of smart environments consider when making decisions about data sharing?
- **RQ2:** How do visitors perceive their privacy in smart environments?
- **RQ3:** What information do visitors need to protect their privacy in smart environments?

To answer these questions, we interviewed 21 participants with a technical background. After inquiring about prior visits of foreign smart environments, we introduce three distinct levels of familiarity with the environment: 1) familiar, 2) known, and 3) new. For each environment, we provide specific use cases to nudge participants to think as visitors of smart environments.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MUM 2020, November 22–25, 2020, Essen, Germany

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-8870-2/20/11…$15.00
https://doi.org/10.1145/3428361.3428464

We show that different environments and contexts impact the data sharing behaviour and intention of visitors more than specific devices. Although our participants had a technical background, they demonstrated difficulties in determining the physical space that is effected by the data collection and judging whether IoT devices capture personal data about them. In general, the privacy preferences of the visitors aligned with those of the owners of IoT devices. However, visitors are limited in making judgements about the data collection and expressing their privacy preferences because they do not have access to adequate information. Based on that, we demonstrate prerequisites that visitors need to control the conditions under which their personal data is captured and processed. We conclude with actionable measures to facilitate this control and motivate future courses of action to improve the control.

2 BACKGROUND AND RELATED WORK

Generally, we refer to privacy as the possibility to control the conditions under which personal data is captured and processed by a third party [12]. This means that each (main) user individually *decides* about these conditions. While the owners of IoT devices have the possibility to exert control, visitors nowadays have very limited courses of action. Privacy perceptions of bystanders have already been studied in different domains. Hence, we detail related work on bystander privacy in related domains, privacy in smart environments and bystander privacy in smart environments.

2.1 Bystander Privacy in Related Domains

Studies of life-logging technologies have shown that bystanders want and need to know whether their actions are about to be recorded in order to give or withdraw their consent [25, 28, 39]. Furthermore, it has been shown that life-logging technologies have to be designed specifically to respect bystanders [21], e.g. by providing obfuscation [3, 17]. These results have also been shown for the scopes of augmented [9, 53] and mixed reality [15, 38, 41], where devices with cameras, such as head-mounted displays, are used. Privacy concerns of bystanders regarding the presence of augmented reality wearable devices can depend on the usage context as two user studies show [9, 16]. In these studies, participants explicitly differentiated between public and private environments. They articulated that they want to be asked for permission before being recorded in private environments.

2.2 Privacy in Smart Environments

IoT devices in smart environments provide their functionality based on access to data about their users and the users' environments collected by sensors. This data access and networking capabilities have resulted in privacy concerns [1, 31, 49, 61]. One stream of research has focused on the concerns of (prospective) IoT device users [2, 5, 7, 8, 10, 19, 54, 55, 57, 59]. When making decisions about their privacy, users differentiate between different environments [19]. Data collection in public is perceived as less critical than in private environments. On the other hand, studies also show that the entity collecting the data is even more important than the environment [32]. This is confirmed by studies showing that IoT device users are furthermore concerned about the data processing by the providers of IoT devices [46, 49]. Based on that, the owners of IoT devices wish to be aware of data that is transferred to device providers [19, 26, 37]. Privacy-related information should be available when buying IoT devices [20]. However, convenience provided by IoT devices is a major factor for their users to sacrifice privacy [19, 60].

2.3 Bystander Privacy in Smart Environments

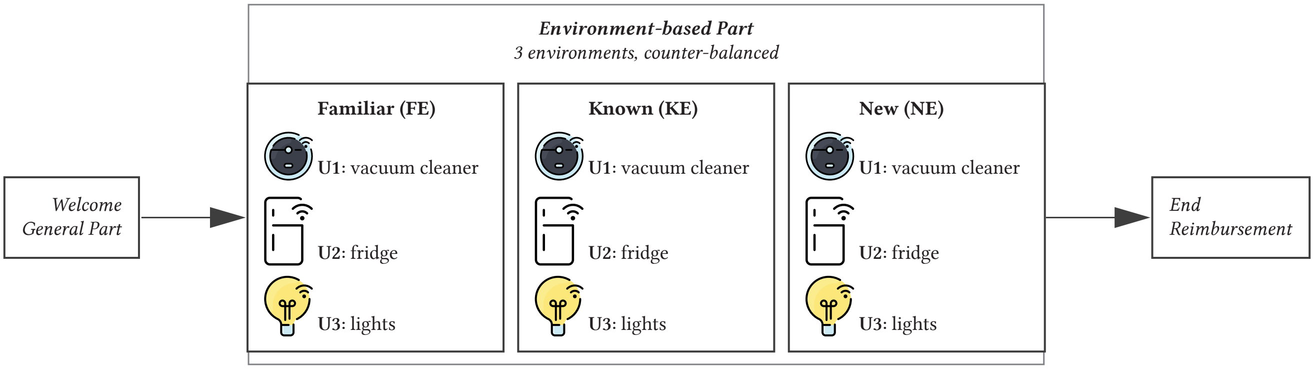
Privacy perceptions regarding smart devices were repeatedly studied in the literature (e.g. [2, 5, 7, 8, 10, 19, 31, 57, 59]). Privacy concerns form a barrier for people in becoming smart home users [1, 49, 55] and having information about the data collection can dismantle this barrier. Hence, (prospective) smart device users wish to be aware of which data is collected by providers [19, 26, 37], ideally prior to purchase [20]. However, smart devices affect not only their owners, but also visitors that are present the environment. The device owners' perspective towards bystanders has been investigated (e.g., [58]) and led to the recommendation of visitor modes that prevent bystanders from accessing the owner's data. An interview study investigating privacy perceptions that results from the presence of bystanders in smart environments concluded that both sides – users and bystanders – have to be considered when designing IoT devices [34]. However, this is particularly challenging due to IoT devices becoming ubiquitous and the increase of the numbers [34].

Yao *et al.* studied three scenarios in which the privacy of bystanders in smart homes can be relevant [56]: 1) renting an apartment with an Internet-connected security camera, 2) the own child playing with a smart toy when visiting friends, and 3) a spouse installing an Amazon Echo in the own shared apartment. They conducted a co-design study with 18 participants and identified factors that impact and mitigate bystanders' concerns in these scenarios. Their most prominent finding was the wish of bystanders to exert control over data collection by specifically interacting with the foreign device. Mare *et al.* specifically investigated IoT devices in AirBnBs [33]. Song *et al.* investigated means to improve the discoverability of IoT devices in smart environments that are private and unknown to visitors, such as hotel rooms [47]. Their results indicate that a combination of LED indicator lights and a beep sound supports users best in discovering IoT devices.

While bystanders can be residents of the owners' household, we especially consider *visitors* a subgroup of bystanders. Shared living and multi-user scenarios have been investigated in the literature already. IoT devices in shared living spaces can lead to tensions between primary and secondary users [31] and discomfort [36, 45, 51], but also to privacy aspects being negotiated over time [23].

2.4 Summary

Adding to the existing body of research, we extend existing results by investigating different contexts of data collection. As contexts, we consider different types of smart environments (i.e., smart environments that are familiar, known and new to visitors). We focus on both, private and public environments. We demonstrate that visitors need to be able to exert control over the conditions under which their personal data is captured and processed. We conclude with actionable measures to facilitate this control.



Icons by Vitaly Gorbachev from www.flaticon.com

Figure 1: Our semi-structured interviews were structured as follows: after a general part, we investigated 3 environments (FE, KE, NE) in counterbalanced order. Within each environment, we explored the 3 same use cases with participants (U1, U2, U3).

3 INTERVIEW STUDY

To investigate the privacy perceptions of a smart environment visitors, we conducted semi-structured interviews with 21 participants. Our choice was motivated by the fact that semi-structured interviews, on the one hand, offer a degree of standardization while, on the other hand, leaving room to investigate the participants' answers in more depth [40]. Before determining the final interview guide, we conducted pilot interviews. Based on those, we adapted interview questions and explanations to improve their clarity. The results from the pilot interviews are not included in this paper.

3.1 Interview Procedure

The procedure of our interviews was as follows (cf. Fig. 1).

3.1.1 Welcome. Before the interview, participants received a consent form with information about the study and data protection policy. After signing it, each participant provided demographics and filled out the ATI-scale to assess affinity for technology [22].

3.1.2 General Part. Interviews started with general questions about the participants' prior experiences with smart devices and smart environments. We asked them which devices they use or have used in the past and whether they already interacted with a device that was owned by another person or entity. Next, we focused on the visit of smart environments. We asked participants about their behaviour in a foreign environment equipped with smart devices.

3.1.3 Environment-Based Part. In the second part of the interview, we gradually introduced three environments for a visit that differed regarding the level of familiarity with the environment. To avoid sequential effects, we varied the order of these environments according to a Latin square [52] for each participant. The investigated environments were as follows:

Familiar Environment (FE) refers to an environment that participants know very well and visit frequently, such as the apartment of a good friend or close relatives.

Known Environment (KE) refers to an environment known by the participants. Participants have visited the environment multiple times but visits are rather rare. This could be a shared work space or waiting area.

New Environment (NE) refers to an environment that participants have never visited before and are likely not to visit again. This could be, for instance, the apartment of a client or a rental home during vacation.

In each environment, we introduced three use cases to nudge the participants to consider a smart environment visit. Hereby, we focused on smart devices that are already available on the market. The investigated use cases were as follows. As first use case we consider a smart vacuum cleaner operating in the room with the visitors (*U1: Vacuum cleaner*). The second use case is a smart fridge re-ordering a product visitors have consumed (*U2: Fridge*). And finally, the third use case is based on automatically controlled smart light bulbs in the room where the visitor is present (*U3: Lights*).

The smart lights represent a use case in which the smart device and the sensors thereof are present in a steady location in the room and function without interaction by the visitors. Once discovered by the visitors, they know the location of the smart device. The fridge also represents a steady location. However, interaction of the visitor is required for the fridge to capture data. Finally, the vacuum cleaner represents a use case in which the smart device and its sensors do not have a steady location. We intentionally did not include camera-based devices and smart voice assistants since those have been investigated in prior studies [56].

For each use case, we considered the following identical questions. First, we let participants explain their understanding of the smart device and its data collection. Next, we asked participants about their behaviour as a visitor. This included the data that they are willing or accepting to share. Since related work has shown that perceived benefits constitute a major factor when consenting data sharing [19, 60], we specifically asked whether the participants would share data in exchange for a benefit. The next question focused on domains within the smart environments in which the participants wish access to information and what kind of information they wish. Finally, we asked for expectations on the providers and manufacturers of the smart devices.

3.1.4 End and Reimbursement. After the interview, the participants were given the opportunity to ask questions. Finally, the examiner reimbursed the participants with a € 10 Amazon voucher.

3.2 Participants and Recruitment

We recruited participants on a university campus in the departments of computer science, business informatics and among students majoring in the combination subject of psychology and information technology, using mailing lists and social networks.

Our participants were on average 25.52 years old (SD = 2.62, $Median$ = 26, Min = 18, Max = 28). One third of the participants identified as female (N = 7), the remainder identified as male (N = 14). We also provided the options "prefer not to say" and "other", but none of our participants chose that. Sixteen participants were university students and five were full-time employees. Our sample demonstrated an average affinity for technology (ATI) [22] of 4.21 (SD = 0.83, $Median$ = 4.33, Min = 2.11, Max = 5.56), where six denotes the highest technical affinity.

3.3 Ethical Considerations

Our institution is located in a country with no requirement for a formal IRB process prior to interview studies. However, the ethics committee at our institution provides a set of guidelines for user studies that we followed. We limited the collection of personal data to a minimal amount to preserve the privacy of our participants. Each participant received a randomly assigned 4-digit identifier. Before the study, each participant received a consent form and information sheet. The information sheet detailed the study's data protection policy and risks associated with the participation. After reading the information sheet and consent form, participants were asked to sign the consent form which was then stored separately from all other captured data. Furthermore, our study, complied with strict national privacy regulations in the authors' country.

Before the analysis, we transcribed all audio recordings into written form and deleted the source files. During transcription, personal information was replaced by neutral placeholders[1].

3.4 Data Analysis and Limitations

First, we transcribed all audio recordings into written form. Next, we analysed the interview transcripts using grounded theory [35]. Our analysis consisted of open and axial coding where two authors of the paper were the coders. The coders individually coded two representative interviews using thematic analysis with open coding [6]. Then, they established a coding tree in a review meeting that consisted of 190 defined codes. To avoid excluding data types and devices mentioned by the participants, the coders established a common code structure used to build new codes. This code structure was used when a participant reported either a new IoT device, a new data type, or new data flow. Each researcher coded one half of the transcripts using the coding tree. During this analysis, 19 new codes were assigned matching the code structure.

The code assignments were discussed in a final review meeting. We concluded with axial coding to relate the codes to each another. This resulted in three levels of codes. In particular, five main categories, namely 1) data sharing, 2) data collection, 3) device providers, and 4) IoT devices, and 5) information, emerged as well as subcategories. For the structure of our coding tree and the individual codes, the reader is referred to Appendix A.

We investigated IoT devices already commercially available. Furthermore, networked smart environments with a variety of devices are as of 2020 still rather scarce. Thus, the presented measures do not consider scalability aspects of smart environments. As a starting point to determine required information and measures, we recruited a rather young tech-savvy sample. This is also reflected by the mean ATI scale of 4.21. Yet, it cannot be assumed that visitors of smart environments have a technical background.

4 RESULTS

In this section, we present the results of our interview study based on the answers to the general questions and the five main categories: data sharing, data collection, information, device providers, and smart devices. Whenever meaningful, we provide comments from participants with a reference to the respective environment, if applicable. In this context, FE denotes familiar environment, KE denotes known environment, and NE refers to new environment. The use cases are abbreviated as follows: U1 refers to the vacuum cleaner, U2 to the fridge, and U3 to the light bulbs (cf. Figure 1).

4.1 General Questions

In the beginning, we asked general questions about IoT devices and smart environments about the participants' experiences, purchases of IoT devices and prior visits of foreign smart environments.

4.1.1 Experiences. All participants were familiar with IoT devices in general and owned at least a smartphone. Nineteen participants had previously interacted with IoT devices owned by others. When asked for the specific IoT devices, participants named voice assistants (e.g., Amazon Echo), smart window shutters, smart light bulbs, smart TVs, smart vacuum cleaners, and smart fridges.

4.1.2 Purchase Decisions. After talking about their experiences with IoT devices, we interviewed them about reasons for (not) purchasing IoT devices. Convenience was the most frequently mentioned reason for owning IoT devices. 17 of our 21 participants stated that smart home devices facilitate daily life by *reducing workload*. The second most frequently mentioned reason (N = 5) was an improved *energy efficiency*. Reasons for not investing in IoT devices were *high cost* of the devices on the market (N = 13) and concerns about *data collection* and *data protection* (N = 9).

4.1.3 Visiting Smart Environments. Furthermore, we asked participants about their general behaviour when visiting a smart environment and whether they would generally alter their behaviour in the presence of IoT devices. One third (N = 7) answered this question affirmatively. One participant even mentioned to visit the home of a friend less frequently due to their IoT devices. Sample statements given by our participants are:

> *"I would try to do less embarrassing things when I'm alone."* P3

> *"It depends on the equipment that's there. If a surveillance camera is hanging somewhere, I think I'd behave differently. I'd think about: what do I want to reveal about myself considering that in the worst-case scenario it ends up somewhere it shouldn't. [names examples]. Regarding audio it's similar [names examples]. With other devices, I don't think I would have such a problem. For a motion sensor for light, I wouldn't change my behaviour."* P4

[1] As an example, the sentence *"my sister Lisa from New York"* would have been transcribed into *"my* [relative placeholder] *from* [city placeholder]*"*.

> *"If they use many of such devices, I would definitely visit them less often. If they had a fridge or a light only, I would be there just as often."* P19

Three participants particularly mentioned they would visit friends with IoT devices more frequently as they are curious.

4.2 Data Sharing

We investigated data sharing behaviour of smart environment visitors. Participants differentiated between *not sharing* specific data, *limited sharing* of specific data, *conditional* sharing (e.g., for personal benefit), and *sharing* of specific data without restriction. In the following, we describe the data that the participants would share without restrictions, and the impact of the familiarity to the environment, its owner and to the IoT devices on data sharing, and on sharing of sensitive data.

4.2.1 Known and Required Data. Eight participants stated to be comfortable with sharing data about them that is *already known* or obvious. This was mentioned in all three investigated environments. Participant 10, for instance, stated:

> *"If he [the doctor] sees that I am present and I can keep quiet while the vacuum cleaner operates [...], the information already exists anyway."* P10, NE, U1

> *"I don't care, my friend knows that I'm using the fridge anyway."* P20, FE, U2

The second kind of data participants were generally comfortable with sharing is data that is *required* in the respective use case or transferred by interaction with the IoT device ($N = 3$):

> *"I'm okay with that. If it only films when I open the fridge and take something out, I think that's okay."* P10, NE, U2

4.2.2 Fear of Consequences. Four participants would refrain from sharing their data, if they fear *negative consequences*, for instance, in the form of long-term behaviour tracking or the occurrence of costs. This was exclusive to known and new environments in which the owner of the IoT device gets information that might harm the participant:

> *"I find employers more critical again, if my employer gets the information, I would be much more careful. I would probably not say anything sensitive, because I simply did not want my employer to hear that."* P10, NE, U3

> *"Let's say I tell you about last weekend on Monday. I don't know if an employer wants to know that I did extreme sports or had a night out. From this, conclusions could be drawn about my lifestyle and I could be classified as a risk employee."* P11, NE, U3

4.2.3 Familiarity. The fear of consequences is also related to the participants' familiarity with the environment, the owner of the device and data practices of the device provider. Missing familiarity with the environment was mentioned by three participants as a reason for not sharing data in known and new environments and further five participants would limited sharing their data. In contrast, eight participants stated that a familiar environment is a reason for them to share data. Hence, the more familiar participants were with an environment, the more comfortable they were with sharing data. Participants, for instance, mentioned:

> *"If have often interacted with these devices, I assume that I know what they do. So I would say that I would share all data without hesitation."* P2, FE, U1

> *"I don't think there is data I would share without hesitation. Especially in a doctor's office or with someone I don't know well, I have to think about what is stored. It could also be that the device records permanently. It first has to be clarified to what extent data is recorded and processed before I would act more freely."* P20, KE

The familiarity to the environment and the environment itself were more important compared to the specific use cases. Seventeen participants only adapted their sharing behaviour to the environment and did not alter it between the use cases. One participant did not differentiate at all, but generally refrained from sharing data.

Furthermore, six participants said their familiarity with the owners of familiar environments and trusting them is a reason for sharing data. One participant stated this about known environments. Another participant stated that data sharing would depend on their trust in the device owner. This indicates that trust in the owner can translate to trust towards the smart environments:

> *"In principle, I trust friends more than I trust an employer - not that I have any quarrel with my employer. But I find it a bit strange if I tell my colleagues something private and then my employer hears it."* P10, KE, U2

> *"My friend knows me well anyways."* P3, FE, U1

Similarly to above, a differentiation between the use cases was not made by the majority of participants, however, the familiarity with the respective smart device was considered:

> *"I don't know exactly whether the refrigerator weighs or whether it has a barcode scanner or a camera or sees who opens the refrigerator and who takes them out. I can't say for sure if it's filming, but if I'm not sure, I assume it films."* P10, FE, U2

4.2.4 Sensitive Data. Even if participants were in familiar environments, 14 participants stated that they would in general not share data that they consider to be sensitive. Among these sensitive data, four participants mentioned bank data, three mentioned health-related data, and one mentioned private pictures. All of them considered these data as too private even in familiar environments:

> *"Certainly I would not give out very sensitive data such as health data or bank account information."* P10, FE

> *"Shared data about whether I'm closing a shutter would not be important to me. For, e.g., video recordings, I would ask again whether this is stored on a server on the Internet or [...] locally and what is done with it, which provider it is."* P7, FE

Four participants would limit the sharing of data that they consider to be sensitive. However, participants were generally comfortable with sharing non-sensitive and non-personal data. Two participants explicitly named small-talk.

> *"It [the fridge] does not know who took the product."* P18, NE, U2

> *"Certainly I would not give out very sensitive data such as health data or account information. There would be no hesitation in*

> *small talk or things I would say at home on the couch where I wouldn't want that Alexa listens to."* P10, FE

Before sharing sensitive data, all participants wanted to explicitly be asked for their consent in any environment and use case.

4.2.5 Sharing for Benefit. We specifically asked participants how they consider the ratio between the sharing of (sensitive) data and gained personal benefits, such as additional comfort. In familiar and known environments, participants were more comfortable to share data if they received a benefit in exchange for that ($N = 15$). In new environments, participants expressed that such a benefit should be very high as otherwise they would refrain from using the device ($N = 3$).

> *"A little comfort for a little sensitive data, I'm okay with that. But, not highly sensitive data [in exchange] for high comfort."* P12, FE

> *"The smart home devices can increase the convenience with use, but for me, that would be too non-secure because I don't know the system and I would not use it due to concerns that too much of my data is being collected."* P19, NE, U2

4.3 Data Collection

The previous category considered reasons for (not) sharing data meaning that sharing is initiated by visitors. In this category, we show participants' acceptance of data collection and their reasons.

4.3.1 Collection of Obvious Data. Similar to the sharing of obvious data detailed above, participants expressed that the collection of obvious data is generally acceptable for them. This, in particular, concerns data that cannot be linked to the participants' identity, such as the presence of a person in the room that can be used for controlling lights or routing the vacuum cleaner. Participants mentioned this throughout all environments and use cases. Sample comments are:

> *"I don't care. A sensor would create data, whether you are in or not, but it would still be okay for me. For turning off lights automatically, [data collection] would be okay."* P16, NE, U3

> *"I don't see any problem there. [...] I don't think the vacuum cleaner collects data that directly concern my person."* P19, KE, U1

4.3.2 Collection of Personal Data. The collection of data that could be linked to the visitors' identity without consent or interaction was perceived as critical within all environments and use cases. Participants, for instance, stated:

> *"I don't know if the refrigerator would know whether I or my friend consumed it [the product]. If the fridge knew that it was me, I would also find that critical."* P17, FE, U2

> *"I would find it odd, particularly in a doctor's office, if video or audio recordings could be made or if there were devices that could do this, because I would also like to talk to my family doctor about things that I do not want to share elsewhere or that are somewhere on the Internet."* P4, KE

4.4 IoT Devices

Considering the specific use cases, we asked participants which entity can access the collected data according to their understanding. Our participants had a technical background and mentioned different locations and possibilities for storing the data. As detailed above, some participants already gained experiences with the IoT devices in the use cases (smart vacuum cleaners, smart lights and smart fridges). Knowledge about these IoT devices did not influence the participants' answers as they were aware of different possibilities to configure these devices. This means that IoT devices with the same functionality from different providers might have different sensors and thus capture different data.

Many participants mentioned that data is stored outside the smart environment, e.g., in a cloud ($N = 19$) or, more specifically, in a storage in a cloud hosted by the device provider ($N = 7$).

When describing the data flow in the individual use cases, participants focused on the functionality. Hence, they mostly described data flow necessary within the use case. Further possibilities, such as using the data to track behaviours, were not detailed.

4.4.1 U1: Smart Vacuum Cleaner. Nine participants stated that data needed for routing the vacuum cleaner, such as the room layout, is only stored on the device and not elsewhere. This data collection was considered to be acceptable for several reasons. Four participants in particular stated that the data is generic presence data that cannot be linked to their identity. Sample comments are:

> *"The vacuum cleaner scans the floor, detects people, looks for dirt and where to clean. The data is then stored in the vacuum cleaner and used for the future."* P8, U1

> *"Well, it scans me and well, now I don't think that it scans the room in full height, all persons, but probably only the area it needs, the floor, if it scans the area in front of it three-dimensionally, if not, then it doesn't matter."* P13, U1

Five participants considered that the vacuum cleaner might not possess the required computational capabilities to calculate the route. Instead, they stated that the vacuum cleaner sends the room layout to a cloud, which answers with a routing plan. These participants considered the collection of these data to be unacceptable if the video could be linked to their identities.

> *"The vacuum cleaner comes in, also into the room, switches on, simply delivers data about the room, objects are measured and perhaps also the dust intensity, and sends it back to the cloud. The cloud determines the stains that need to be cleaned more compared to last time. It then sends the route plan back to the vacuum cleaner and it drives off. "* P2, U1

> *"The vacuum cleaner collects the data, that there is someone and the data where does it flow to? They are logically processed by it and finally probably put back into some cloud based memory, whatever, where they are evaluated. From there, the user or the owner of the robot can probably access it again and evaluate the data."* P14, U1

4.4.2 U2: Smart Fridge. Considering the fridge, nine participants stated that the content or the order are processed online, which is necessary for the provided functionality. Similar to the vacuum cleaner, the collection of this data was considered acceptable if not linked to their identities by the owners of the smart fridge or the provider. Furthermore, participants feared a loss of control regarding the orders of the fridge in new environments. Sample comments by participants are:

> *"I access the system, the refrigerator. This is a physical interaction. Then the refrigerator creates the data, collects it, processes it and sends it to another unit or sends the command, 'reorder'. It receives an order confirmation. At the moment it is delivered and put back in, the refrigerator again collects the data and knows that the target stock is reached again. "* P1, U2

> *"We have another problem, and that is a much more intimate behaviour in an AirBnb-apartment, because of that it is of course problematic if the owner of the smart fridge - in this case the owner of the AirBnb-home - notices what kind of food I consume there and automatically reorders it. It takes over the control of what I want to consume in the refrigerator during the time I am there. "* P20, NE, U2

4.4.3 U3: Smart Light Bulbs. Finally, 16 participants stated that the status of the light bulbs is available via the Internet to enable the functionality of remote control. These data was also considered to be acceptable if it cannot be linked to the visitor's identity. However, one participant stated that if it is known to someone that a visitor is present it could be determined whether this visitor has left a place based on the status of the lights. This participant was reluctant to share even generic data about their presence.

> *"The lights should just give the signal: I'm still on and it's Sunday afternoon. Then he could just turn that off. On the other hand, the sensors could also determine whether someone is still in the office or not. "* P6, KE, U3

> *"It depends a bit on whether I can operate the equipment. If I go into [my friend's] apartment and can only turn the lights on when I operate the tablet, but the tablet records who opened it because it is locked and I used a guest access point, then maybe he [my friend] knows I turned the lights on. But if I had simply turned it on with the light switch and he turned it off again and had not installed a camera to check - which I have not assumed I did with the lamps - I would say there was no information."* P10, KE, U3

4.5 Information

We asked participants if they wished access to information about smart environments and IoT devices.

The majority of participants ($N = 19$) stated to wish information about the *data collection* by the devices in all environments and use cases. They placed a focus on new environments and environments in which a data collection is unexpected. Even if an environment is known or familiar, the IoT devices deployed in it might change. Hence, information about new IoT devices should be made available:

> *"It should be recognisable, which smart home devices are present, and for 'what' they are there and where the data are stored or if data is stored."* P5, NE, U2

> *"Information should be available in any case, especially in public places. [...] I would like to know in any case: 'What kind of devices are there? How do they work? Where is the data stored? But I would leave it up to the individual, whether they are interested in that information. If medical assistants in a doctor's office provide such information and visitor perhaps get an information sheet, then that is enough for me. I can still decide: Do I want to read it or not? Does that interest me? Do I care? I can decide."* P4, KE

When asked to explain the purpose of the information, participants stated transparency ($N = 19$), an establishment of trust ($N = 9$), or personal interest ($N = 8$):

> *"Ultimately, any manufacturer of such devices should clearly state where the data will end up, where the server is located on which the data will end up. That you just have the information when you want it. It should be provided and it should be honest and transparent."* P4, FE

> *"So for me trust, i.e. creating transparency, at the same time... what else could you expect from them? [names examples] I would simply say to create more transparency, to show further foresight. Yes."* P14, KE

> *"This is difficult to say, because I do not constantly deal with it and I know which information is accessible in principle. But yes, it is a bit frightening that I do not know: what exactly is the data that is being collected by the devices? Where does it flow to? How can they be linked? Does it flow via the Internet or directly there? It would be exciting, but it is also a little like utopia."* P10, FE

The information should contain the data that is collected, the physical space that is affected by the collection, and the purpose of collection:

> *"I think, especially in areas where you are often together with strangers, such as in a doctor's office, it should definitely be made clear how the data is processed. Obviously I cannot demand from a stranger that he hangs up a sign in his apartment on which is written: 'We process data in the following way.' However, in all public places, departments and doctors' offices, one expects information on how the data is processed."* P20, KE

Two participants stressed that this information should be prominent in spaces where data collection is unexpected even if the environment is familiar:

> *"I think you should know when you're being recorded somewhere. It should not look like a vase, but like a video camera."* P7, FE

Participants were also aware about the practicability and scalability issues in smart environments with many devices. Two participants, for instance, stated:

> *"It's just difficult. You just can't know about everything. As an indirect user, I would of course wish for that, but I see difficulties in implementing it in everyday life."* P19, FE

> *"I think it [labels] is a good idea but I think it's difficult to implement, because then, probably in every larger office building all warning signs would have to go: 'In the next 5m2 is a microphone, which could record your voice', and every subway station: 'Attention, video is being recorded here'. Then you have such a flood of information that you don't care. [...] I think that's very, very difficult to implement."* P17, KE

4.6 Device Providers

We also asked the participants about their expectations towards the providers.

4.6.1 Transparency. More transparency about data practices was mentioned 14 times throughout all environments and use cases. The data practices entailed data security-related information ($N = 9$), sharing with other entities ($N = 7$), and storage on the provider's servers ($N = 2$). Sample statements from our participants are:

> *"In any case, it's transparency. I know it is not an easy subject, because how do you intend to create transparency? That they [the providers] need the data somewhere and use it anonymously is something I find perfectly acceptable. But I just don't know if they [the providers] will indeed do that. I also find it extremely difficult to trace."* P11, KE

4.6.2 Control. Five participants also expressed that providers should enable better control about data shared and collected by the devices. This was justified by a wish for improved self-determination:

> *"I would like to know about all data and recordings that are collected about me. Instead of only storing and analysing it somewhere, I also want to express 'No, I don't want it to be stored.' And I also want to be sure that when I say that I want it to be deleted, it will indeed be deleted. There is always a background fear that it may still be on some of the company's servers."* P19, FE

4.7 Results Summary

In summary, participants differentiated the IoT devices based on the provided functionality. In all use cases, they considered data collection to be generally acceptable by a device if it cannot be linked to their identity. If such a link can be made, participants consider their familiarity to the environment, to the device owner and the sensitivity of the shared data. Considering new environments, participants feared a loss of control, because an IoT device might act unintendedly. Considering information about data collection, our participants asked for information in domains where data collection is unexpected. They furthermore expressed to judge the capabilities of the IoT devices and wished information about that. Device providers should be transparent with regards to their data practices and security measures as a basis to control what happens to their data.

5 DISCUSSION

In this section, we first discuss the *results from our interviews* based on the categories that emerged in the code analysis. Then, we detail *prerequisites for informed privacy decisions* that visitors of smart environments need. Based on these prerequisites, we finally discuss *possible measures* that could be applied in smart environments (cf. Figure 2) and their limitations.

5.1 Interview Results

In this section, we discuss the results from our interview study.

5.1.1 Data Sharing. We investigated three levels of familiarity to smart environments. Participants adjusted their data sharing behaviour to the respective smart environment, IoT devices, and the owner of the device or environment. If the participants knew the specific IoT device and its capabilities, they adjusted their sharing behaviour to the device. If the device was unknown, the behaviour was adapted to the familiarity of the environment. The owner of the environment played an important role because they might have access to potential privacy-sensitive data. In contrast participants only rarely differentiated between our three investigated use cases. These results have already been demonstrated in investigations of IoT device owners [46, 49].

Furthermore, participants differentiated whether the smart environment was a public or private environment when making decisions about sharing their data. In general, in familiar and private environments, participants were most conformable with sharing (sensitive) data. In contrast, they were least comfortable with data sharing in new and public environments. This shows intervention and transparency methods to be most crucial in public smart environments. This also mirrors results from previous studies of IoT device owners [19] and extends them to the visitors' perspective.

Our results show that visitors and owners of smart environments make similar considerations about sharing their data. While owners of smart environments are typically aware of IoT devices being present, it cannot be assured that visitors are likewise aware. This stresses the importance of methods to raise visitors' awareness.

5.1.2 Data Collection. Considering data collection, visitors are more comfortable with sharing obvious data or data that is transferred by an intended interaction with the IoT device. This also extends results of previous studies [19] to the scope of visitors.

20 of 21 participants would still visit friends owning IoT devices, even if they were uncomfortable with data being collected. This shows that social aspects are valued stronger when making privacy decisions. If visitors are not equipped with adequate knowledge and measures to protect their privacy, they are likely to sacrifice their privacy in private smart environments, as results indicate.

5.1.3 IoT Devices, Information, and Device Providers. Our sample consisted of young, rather tech-savvy adults. When asked about the data flow of different devices they were aware that different device providers might implement different data flows. Based on that, several participants assumed a worst-case scenario, i.e. a data flow that penetrates their privacy most. Since even our rather tech-savvy participants struggled in determining the correct data flow, visitors that are less tech-savvy might struggle even more.

When asked for information about data collection, participants wished to generally be aware about it. This is also connected to results about the owners of IoT devices [19, 26, 37] and bystanders [34, 56]. Participants furthermore wished for information about the purposes of data collection, especially in environments in which this is unexpected. This was prominent in known and new environments in which visitors are most vulnerable since IoT devices might be very discreet.

In the following, we demonstrate prerequisites that visitors need to enforce an informed decision about their privacy. This is followed by specific possibilities to implement measures.

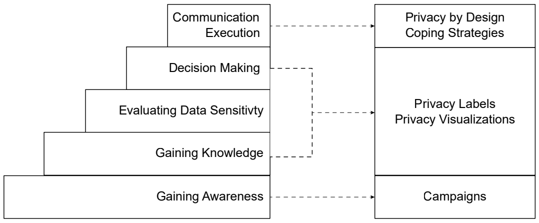

Communication Execution
Decision Making
Evaluating Data Sensitivty
Gaining Knowledge
Gaining Awareness
Privacy by Design Coping Strategies
Privacy Labels Privacy Visualizations
Campaigns

Figure 2: Prerequisites for visitors (left) for making and communicating privacy decisions and possible measures (right).

5.2 Prerequisites for Informed Privacy Decisions for Visitors

Privacy for visitors describes their ability to control the conditions under which their personal data is captured and processed in a smart environment. This means that visitors have to be able to gather privacy-related information in order to make a decision. Next, we detail consecutive steps building upon each other to reach such a decision.

5.2.1 Gaining Awareness. As a foundation, visitors need to be aware that the data collection in a smart environment can impact their privacy. A possible barrier is the misconception that the data of visitors cannot be privacy-sensitive due to the scarcity of visits as our investigation shows. Furthermore, a visitor might over-trust the environment or base their trust only on the device owner. Another aspect that impacts awareness is given by the fact that the increasing number of IoT devices and the different configuration possibilities constitutes a challenge [34].

5.2.2 Gaining Knowledge. Second, visitors need information about the data collection. Participants in our and related studies [34, 56] wished information about the data that is collected, the physical space that is affected by the collection, and the purpose of collection. This is particularly crucial in environments that visitors visit only rarely or just once, but also in familiar environments in which IoT devices may change.

5.2.3 Evaluating Data Sensitivity. In this stage, visitors evaluate whether they consider the collected data to be privacy-sensitive. Participants found it acceptable to share obvious data or data based on an intended interaction. In case data is not considered sensitive, no further actions are needed. A possible barrier in this step might be that the visitors misjudge the impact of sharing their data. For instance, it has been demonstrated in other domains that people are comfortable sharing their location as it is not connectable to their person, but that location data can be de-anonymized [24].

5.2.4 Decision Making. If visitors are aware about impacts on their privacy, have knowledge about the data collection, and consider the data to be sensitive, they can make a decision whether they want their data to be captured. If information about the specific devices is unavailable, this decision can still be informed by the familiarity to the smart environment, its owner, purpose and context.

5.2.5 Decision Communication and Execution. Once the decision is made, visitors have to be able to communicate it to the affected IoT devices in the smart environment. The smart environment could either react to the visitor's decision, or the visitor could perform an action to enforce it.

After those five steps, visitors have made an informed privacy decision and communicated it to their environment. Another possible exit point is step three, after which no decision communication or further measures are required.

5.3 Possible Measures and Future Work

In the following, we provide concrete examples demonstrating how to support the aforementioned steps and ultimately enable privacy protection of visitors in smart environments. Furthermore, we provide opportunities for future work.

5.3.1 General Awareness. As a first step, visitors need to be *aware* of potential privacy violations. This is a general aspect that is independent from the specific smart environment. Thus, methods can be outside a specific smart environment. Methods for gaining awareness can be TV spots or poster advertisements. For a detailed list of such awareness methods, the reader is referred to [44].

5.3.2 Privacy Labels and Visualisations. To further support users to gain awareness, deepen their knowledge, and evaluate data sensitivity, privacy labels and visualisations are viable solutions that can be implemented without altering the IoT devices themselves.

Prior works suggested *privacy labels* as a method for providing concise information, e.g. [18, 20, 27]. Today, such labels have already been introduced in the form of signs by several countries to inform about CCTV in public places. A prior study of bystander privacy in smart homes also suggested distributing physical signs with QR codes providing further information by the device manufacturer that is also accessible for bystanders [56]. Naeini *et al.* suggested labels on the smart device's packing to inform purchase decisions [20]. Suggested content of such labels includes the type of collected data and the frequency of data sharing [18]. However, conventional labels on the device's package are not applicable to visitors as they usually are not involved in the purchase and unwrapping of the product. Also QR codes have limitations. First, they might be overlooked by visitors. Second, the information the QR code links to might not be read or understood completely. Third, it is questionable whether the owners of IoT devices would indeed distribute such QR code stickers in their private homes. Thus, QR code stickers might be a viable solution for public environments if designed carefully. Finally, even if an IoT device would be labelled perfectly, visitors might struggle to determine where exactly within the smart environment their data is collected and in which way it is processed, stored and potentially shared (e.g., with the device manufacturer). This issue has been frequently mentioned by our participants. Even though they had a rather high affinity for technology, they were not able to judge potential tracking spaces accurately. In line with our participants wishing for details on data collection, we suggest labelling in such a way that it is accessible for visitors. At the same time, enforcing labels in private home environments might also prevent owners from deploying devices based on aesthetic or social reasons.

In professional or public environments, this could be solved by making such labels accessible via, e.g. signs that are publicly readable before entering the physical space that is affected by the data collection. Furthermore, such labels should be formulated in easy language and clearly indicate where the device is located, where data is stored and whether it is shared with other entities.

While private spaces might, on one hand, be considered less critical (i.e., participants were more willing to share data in known and familiar environments), it is even more crucial to address needs for detailed information on data collection and sharing. As described above, QR code stickers and labels are unlikely to provide the necessary information. This shows that visitors require active privacy assistance, for instance by a software or IoT device. Several possibilities to assist the owners of IoT devices have been proposed in the literature (cf. [11, 13, 14, 30]). Based on our results, we argue that exploring such assisting methods for visitors and privacy implications for them forms an integral part of future work.

Many participants wished for information about potential data collection, especially in scenarios where they do not expect their personal data to be collected. This indicates a need for additionally *visualising* data collecting sources to increase visitors' *awareness*.

Limited means for mode transparency of IoT devices already exist in state-of-the art products and were previously investigated. An example is status indicator lights like small LEDs. A prominent example is webcams. However, users and especially visitors might, on one hand, overlook this feature [43], while on the other hand the indicator might be unreliable. For instance, the software drivers could be manipulated in such a way that a camera is recording while not correctly indicating its actual state. Thus, some camera manufacturers introduced lens covers. More obvious design alternatives to status indicator lights have been suggested in prior work [29, 36, 47]. Further IoT devices indicating their current state include smart speakers (e.g., Amazon's Alexa), where a light ring is changing from red to green while recording [36]. Such alternatives form an intuitive visual indication to determine the device status. Related work suggested additional means such as visual or auditory cues, or contextual pictures providing information on the location of IoT devices [47]. While such solutions might be suitable for new environments (hotel rooms, holiday homes), the deployment in private homes or shared spaces needs further investigation.

Up to now, we have focused on status indication that is integrated in the IoT device. However, visitors who are not used to specific devices or those who might even be unaware of such devices being present, might easily overlook such means of status indication. Furthermore, device-based status indication requires an alternation of the smart device which is time-consuming and visitors might want measures that are more immediately applicable. We thus suggest to employ additional visualisations from a bystander's perspective, e.g. by means of augmented reality (AR). In particular, spaces in which data is collected could be highlighted in AR. For instance, the physical space in which a microphone could capture audio could be visually indicated. Based on the visitor's personal consideration of data sensitivity, the highlighting could be personalised.

The measures described above enable the first three steps, namely gaining awareness and knowledge and evaluating data sensitivity, that are necessary to make informed privacy decisions. We proceed by describing methods for decision communication.

5.3.3 *Empowering Devices to Protect Visitors' Privacy.*

As participants mentioned to behave differently or even stop visits when being tracked in smart environments, admins of smart environments could temporarily deactivate potential data collection to accommodate their visitors. As an example, one of our use cases included a smart fridge. While its smart functionality may provide comfort to the owner, the fridge itself is still performing its duties (i.e., refrigerating groceries) without the "smart" features, such as a camera. To prevent the visitor's privacy to be violated when using the host's kitchen, the fridge's camera could temporarily be turned off. Related work also revealed that residents of smart environments wish for specific visitor modes [34, 58]. Such a mode could be designed in such a way that the device's primary functionality is preserved while the data collection is (temporarily) switched off.

5.3.4 *Empowering Visitors to Protect their own Privacy.*

To ultimately communicate their privacy decision, visitors as of now have limited possibilities. While one possibility is avoidance of smart environments in general, we rather suggest to provide suitable means as coping strategies for visitors to (temporarily) adjust the space's settings to their needs.

We first detail measures that have been suggested in related domains. As such, bystanders could negotiate with the owner of the smart devices to switch it off [23]. This, however, might result in a tension between the owner and this visitor and is time-consuming for a large number of devices. Thus, visitors might sacrifice their privacy due to the effort. Bystanders could furthermore adapt their behaviour or use tools that support them in protecting their privacy. Related work suggested bystanders to carry active noise to prevent audio recordings [53]. Edward Snowden has covered himself with a blanket which can be seen in the documentary Citizenfour[2]. Both examples show how extreme such solutions can get. Using the blanket is likely to disrupt interaction with others and active noise might result in unintended side effects on other devices. A study of bystander coping methods confirmed that bystanders in smart environments nowadays only have limited courses of action which are considered as highly artificial or hypothetical [34].

Additional means to make privacy "graspable" could be applied as well. This refers to physical devices that can be placed on IoT devices for controlling their data collection behaviour. For instance, the privacy hat for Alexa [50] nudges users to use the device's mute button. Further means to permanently hand over the control of privacy settings to visitors should be subject to future work.

6 CONCLUSION

The share of smart devices in households is growing. Data capturing and processing by the devices concern *any* person that is present in the smart home environment, also visitors. In this paper, we presented an interview study to find out prerequisites that visitors need to be able to reach an informed decision about their privacy. We also show that visitors consider their familiarity with the environment and the device when deciding whether they want to share their data. Based on our results, we extract five steps as prerequisites for visitors to exert control over their privacy. Those five steps are: 1) gaining awareness, 2) gaining knowledge, 3) evaluating

[2] https://www.imdb.com/title/tt4044364/, last accessed 16/08/2020

data sensitivity, 4) decision making, and 5) decision communication. For each step, we present specific implementation possibilities. Those possibilities can either be used immediately, such as privacy labels, or motivate future work, such as means for visitors to use augmented reality as a method to visualise data collection.

ACKNOWLEDGMENTS

We would like to thank Andreas Schütz for supporting the data collection within the interview study. This work has been co-funded by the German Federal Ministry of Education and Research (BMBF) within the SWC 2.0 "PrivacyGate" 01|S17050, by the BMBF and the Hessen State Ministry for Higher Education, Research and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE, and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – grant numbers 251805230/GRK 2050, 316457582, and 425869382. The icons used in Figure 1 are by Vitaly Gorbachev from www.flaticon.com

REFERENCES

[1] Noura Abdi, Kopo M. Ramokapane, and Jose M. Such. 2019. More than Smart Speakers: Security and Privacy Perceptions of Smart Home Personal Assistants. In *Proceedings of the Symposium on Usable Privacy and Security (SOUPS '19)*. USENIX Association, Berkeley, CA, USA, 1–16.

[2] Tousif Ahmed, Roberto Hoyle, Patrick Shaffer, Kay Connelly, David Crandall, and Apu Kapadia. 2017. Understanding Physical Safety, Security, and Privacy Concerns of People with Visual Impairments. *IEEE Internet Computing* 21, 3 (May/June 2017), 56–63. https://doi.org/10.1109/MIC.2017.77

[3] Rawan Alharbi, Mariam Tolba, Lucia C. Petito, Josiah Hester, and Nabil Alshurafa. 2019. To Mask or Not to Mask? Balancing Privacy with Visual Confirmation Utility in Activity-Oriented Wearable Cameras. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 3, Article 72 (Sept. 2019), 29 pages. https://doi.org/10.1145/3351230

[4] Florian Alt and Emanuel von Zezschwitz (Eds.). 2019. Special Issue: Emerging Trends in Usable Security and Privacy. *Journal of Interactive Media (icom)* 18, 3 (Dec. 2019).

[5] Noah Apthorpe, Yan Shvartzshnaider, Arunesh Mathur, Dillon Reisman, and Nick Feamster. 2018. Discovering Smart Home Internet of Things Privacy Norms Using Contextual Integrity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (2018), 59. https://doi.org/10.1145/3214262

[6] Ann Blandford, Dominic Furniss, and Stephann Makri. 2016. *Qualitative HCI Research: Going Behind the Scenes*. Vol. 9. Morgan & Claypool Publishers. 1–115 pages.

[7] Denys Brand, Florence D. DiGennaro Reed, Mariah D. Morley, Tyler G. Erath, and Matthew D. Novak. 2019. A Survey Assessing Privacy Concerns of Smart-Home Services Provided to Individuals with Disabilities. *Behavior Analysis in Practice* (2019), 1–11. https://doi.org/10.1007/s40617-018-00329-y

[8] Eun Kyoung Choe, Sunny Consolvo, Jaeyeon Jung, Beverly Harrison, Shwetak N. Patel, and Julie A. Kientz. 2012. Investigating Receptiveness to Sensing and Inference in the Home Using Sensor Proxies. In *Proceedings of the Conference on Ubiquitous Computing (UbiComp '12)*. ACM, New York, NY, USA, 61–70. https://doi.org/10.1145/2370216.2370226

[9] Soumyadeb Chowdhury, Md Sadek Ferdous, and Joemon M. Jose. 2016. Bystander Privacy in Lifelogging. In *Proceedings of the International BCS Human Computer Interaction Conference: Companion Volume* (Poole, United Kingdom) *(HCI '16)*. BCS Learning & Development Ltd., Swindon, UK, Article 15, 3 pages. https://doi.org/10.14236/ewic/HCI2016.62

[10] Hyunji Chung, Michaela Iorga, Jeffrey Voas, and Sangjin Lee. 2017. Alexa, Can I Trust You? *Computer* 50, 9 (2017), 100–104. https://doi.org/10.1109/MC.2017.3571053

[11] Jessica Colnago, Yuanyuan Feng, Tharangini Palanivel, Sarah Pearman, Megan Ung, Alessandro Acquisti, Lorrie Faith Cranor, and Norman Sadeh. 2020. Informing the Design of a Personalized Privacy Assistant for the Internet of Things. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376389

[12] Mary J. Culnan and Pamela K. Armstrong. 1999. Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation. *Organization Science* 10, 1 (1999), 104–115. https://doi.org/10.1287/orsc.10.1.104

[13] A. Das, M. Degeling, D. Smullen, and N. Sadeh. 2018. Personalized Privacy Assistants for the Internet of Things: Providing Users with Notice and Choice. *IEEE Pervasive Computing* 17, 3 (Jul 2018), 35–46. https://doi.org/10.1109/MPRV.2018.03367733

[14] Anupam Das, Martin Degeling, Daniel Smullen, and Norman Sadeh. 2018. Personalized Privacy Assistants for the Internet of Things: Providing Users with Notice and Choice. *IEEE Pervasive Computing* 17, 3 (Jul 2018), 35–46. https://doi.org/10.1109/MPRV.2018.03367733

[15] Jaybie A. de Guzman, Kanchana Thilakarathna, and Aruna Seneviratne. 2018. Security and Privacy Approaches in Mixed Reality: A Literature Survey. Cryptology ePrint Archive, Report 1802.05797. (2018), 1–40. https://doi.org/10.1145/3359626 https://arxiv.org/pdf/1802.05797.pdf.

[16] Tamara Denning, Zakariya Dehlawi, and Tadayoshi Kohno. 2014. In Situ With Bystanders of Augmented Reality Glasses: Perspectives on Recording and Privacy-Mediating Technologies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, 2377–2386. https://doi.org/10.1145/2556288.2557352

[17] Mariella Dimiccoli, Juan Marín, and Edison Thomaz. 2018. Mitigating Bystander Privacy Concerns in Egocentric Activity Recognition with Deep Learning and Intentional Image Degradation. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4, Article 132 (Jan. 2018), 18 pages. https://doi.org/10.1145/3161190

[18] Pardis Emami-Naeini, Yuvraj Agarwal, Lorrie Faith Cranor, and Hanan Hibshi. 2020. Ask the Experts: What Should Be on an IoT Privacy and Security Label? arXiv:cs.CY/2002.04631

[19] Pardis Emami-Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Cranor, and Norman Sadeh. 2017. Privacy Expectations and Preferences in an IoT World. In *Proceedings of the Symposium on Usable Privacy and Security (SOUPS '17)*. USENIX Association, Berkeley, CA, USA, 399–412.

[20] Pardis Emami-Naeini, Henry Dixon, Yuvraj Agarwal, and Lorrie Faith Cranor. 2019. Exploring How Privacy and Security Factor into IoT Device Purchase Behavior. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. ACM, New York, NY, USA, Article 534, 12 pages. https://doi.org/10.1145/3290605.3300764

[21] Md Sadek Ferdous, Soumyadeb Chowdhury, and Joemon M. Jose. 2017. Analysing privacy in visual lifelogging. *Pervasive and Mobile Computing* 40 (2017), 430–449. https://doi.org/10.1016/j.pmcj.2017.03.003

[22] Thomas Franke, Christiane Attig, and Daniel Wessel. 2019. A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale. *International Journal of Human-Computer Interaction* 35, 6 (2019), 456–467. https://doi.org/10.1080/10447318.2018.1456150 arXiv:https://doi.org/10.1080/10447318.2018.1456150

[23] Christine Geeng and Franziska Roesner. 2019. Who's In Control?: Interactions In Multi-User Smart Homes. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, 268. https://doi.org/10.1145/3290605.3300498

[24] Philippe Golle and Kurt Partridge. 2009. On the Anonymity of Home/Work Location Pairs. In *Proceedings of the 7th International Conference on Pervasive Computing* (Nara, Japan) *(Pervasive '09)*. Springer-Verlag, Berlin, Heidelberg, 390–397. https://doi.org/10.1007/978-3-642-01516-8_26

[25] Roberto Hoyle, Robert Templeman, Steven Armes, Denise Anthony, David Crandall, and Apu Kapadia. 2014. Privacy Behaviors of Lifeloggers Using Wearable Cameras. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Seattle, Washington) *(UbiComp '14)*. ACM, New York, NY, USA, 571–582. https://doi.org/10.1145/2632048.2632079

[26] Timo Jakobi, Corinna Ogonowski, Nico Castelli, Gunnar Stevens, and Volker Wulf. 2017. The Catch(Es) with Smart Home: Experiences of a Living Lab Field Study. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. ACM, New York, NY, USA, 1620–1633. https://doi.org/10.1145/3025453.3025799

[27] Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W Reeder. 2009. A" nutrition label" for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security*. 1–12.

[28] Marion Koelle, Matthias Kranz, and Andreas Möller. 2015. Don'T Look at Me That Way!: Understanding User Attitudes Towards Data Glasses Usage. In *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services* (Copenhagen, Denmark) *(MobileHCI '15)*. ACM, New York, NY, USA, 362–372. https://doi.org/10.1145/2785830.2785842

[29] Marion Koelle, Katrin Wolf, and Susanne Boll. 2018. Beyond LED Status Lights-Design Requirements of Privacy Notices for Body-Worn Cameras. In *Proceedings of the International Conference on Tangible, Embedded, and Embodied Interaction (TEI '18)*. ACM, New York, NY, USA, 177–187. https://doi.org/10.1145/3173225.3173234

[30] Marc Langheinrich. 2002. A privacy awareness system for ubiquitous computing environments. In *international conference on Ubiquitous Computing*. Springer, 237–245.

[31] Josephine Lau, Benjamin Zimmerman, and Florian Schaub. 2018. Alexa, Are You Listening?: Privacy Perceptions, Concerns and Privacy-Seeking Behaviors With Smart Speakers. *Proceedings of the ACM Conference on Human-Computer Interaction* 2, CSCW (2018), 102. https://doi.org/10.1145/3274371

[32] Scott Lederer, Jennifer Mankoff, and Anind K. Dey. 2003. Who Wants to Know What When? Privacy Preference Determinants in Ubiquitous Computing. In *CHI*

'03 Extended Abstracts on Human Factors in Computing Systems (Ft. Lauderdale, Florida, USA) *(CHI EA '03)*. Association for Computing Machinery, New York, NY, USA, 724–725. https://doi.org/10.1145/765891.765952

[33] Shrirang Mare, Franziska Roesner, and Tadayoshi Kohno. 2020. Smart Devices in Airbnbs: Considering Privacy and Security for both Guests and Hosts. *Proceedings on Privacy Enhancing Technologies* 2020, 2 (2020), 436–458.

[34] Karola Marky, Alexandra Voit, Alina Stöver, Kai Kunze, Svenja Schröder, and Max Mühläuser. 2020. "I don't know how to protect myself": Understanding Privacy Perceptions Resulting from the Presence of Bystanders in Smart Environments. In *Proceedings of the NordiCHI Nordic conference on Human-computer Interaction (NordiCHI '20)*. ACM, New York, USA.

[35] Terence V. McCann and Eileen Clark. 2003. Grounded Theory in Nursing Research: Part 1 – Methodology. (2003).

[36] Emily McReynolds, Sarah Hubbard, Timothy Lau, Aditya Saraf, Maya Cakmak, and Franziska Roesner. 2017. Toys That Listen: A Study of Parents, Children, and Internet-Connected Toys. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5197–5207. https://doi.org/10.1145/3025453.3025735

[37] Mateusz Mikusz, Steven Houben, Nigel Davies, Klaus Moessner, and Marc Langheinrich. 2018. Raising Awareness of IoT Sensor Deployments. In *Proceedings of the Living in the Internet of Things: Cybersecurity of the IoT*. IET. https://doi.org/10.1049/cp.2018.0009

[38] Vivian Genaro Motti and Kelly Caine. 2015. Users' Privacy Concerns About Wearables. In *Proceedings of the International Conference on Financial Cryptography and Data Security (FC '15)*. Springer, 231–244. https://doi.org/10.1007/978-3-662-48051-9_17

[39] David H. Nguyen, Alfred Kobsa, and Gillian R. Hayes. 2008. An Empirical Investigation of Concerns of Everyday Tracking and Recording Technologies. In *Proceedings of the International Conference on Ubiquitous Computing* (Seoul, Korea) *(UbiComp '08)*. Association for Computing Machinery, New York, NY, USA, 182–191. https://doi.org/10.1145/1409635.1409661

[40] Briony J. Oates. 2005. *Researching Information Systems and Computing*. Sage.

[41] Alfredo Perez, Sherali Zeadally, Luis Matos Garcia, Jaouad Mouloud, and Scott Griffith. 2018. FacePET: Enhancing Bystanders' Facial Privacy with Smart Wearables/Internet of Things. *Electronics* 7, 12 (2018), 379. https://doi.org/10.3390/electronics7120379

[42] Sarah Pidcock, Rob Smits, Urs Hengartner, and Ian Goldberg. 2011. Notisense: An Urban Sensing Notification System to Improve Bystander Privacy. In *Proceedings of the International Workshop Sensing Applications on Mobile Phones (PhoneSense '11)*. 1–5.

[43] Rebecca S Portnoff, Linda N Lee, Serge Egelman, Pratyush Mishra, Derek Leung, and David Wagner. 2015. Somebody's watching me? assessing the effectiveness of webcam indicator lights. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 1649–1658.

[44] Stefanie Pötzsch. 2008. Privacy awareness: A means to solve the privacy paradox?. In *IFIP Summer School on the Future of Identity in the Information Society*. Springer, 226–236.

[45] Olivia K. Richards. 2019. Family-Centered Exploration of the Benefits and Burdens of Digital Home Assistants. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '19)*. ACM, SRC11. https://doi.org/10.1145/3290607.3308458

[46] Tom A. Rodden, Joel E. Fischer, Nadia Pantidi, Khaled Bachour, and Stuart Moran. 2013. At Home with Agents: Exploring Attitudes Towards Future Smart Energy Infrastructures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. ACM, New York, NY, USA, 1173–1182. https://doi.org/10.1145/2470654.2466152

[47] Yunpeng Song, Yun Huang, Zhongmin Cai, and Jason I. Hong. 2020. I'm All Eyes and Ears: Exploring Effective Locators for Privacy Awareness in IoT Scenarios. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376585

[48] Statista. 2019. Smart Home Worldwirde. https://www.statista.com/outlook/279/100/smart-home/worldwide (Accessed January 2020).

[49] Madiha Tabassum, Tomasz Kosinski, and Heather Richter Lipford. 2019. I Don't Own the Data": End User Perceptions of Smart Home Device Data Practices and Risks. In *Proceedings of the Fifteenth Symposium on Usable Privacy and Security (SOUPS '19)*.

[50] Christian Tiefenau, Maximilian Häring, Eva Gerlitz, and Emanuel von Zezschwitz. 2019. Making Privacy Graspable: Can we Nudge Users to use Privacy Enhancing Techniques? arXiv:cs.HC/1911.07701

[51] Blase Ur, Jaeyeon Jung, and Stuart Schechter. 2014. Intruders Versus Intrusiveness: Teens' and Parents' Perspectives on Home-entryway Surveillance. In *Proceedings of the International Joint Conference on Pervasive and Ubiquitous Computing* (Seattle, Washington) *(UbiComp '14)*. ACM, New York, NY, USA, 129–139. https://doi.org/10.1145/2632048.2632107

[52] E. J. Williams. 1949. Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Chemistry* 2, 2 (1949), 149–168.

[53] Katrin Wolf, Karola Marky, and Markus Funk. 2018. We should start thinking about Privacy Implications of Sonic Input in Everyday Augmented Reality! *Mensch und Computer 2018-Workshopband* (2018). https://doi.org/10.18420/muc2018-ws07-0466

[54] Peter Worthy, Ben Matthews, and Stephen Viller. 2016. Trust Me: Doubts and Concerns Living With the Internet of Things. In *Proceedings of the ACM Conference on Designing Interactive Systems (DIS '16)*. ACM, New York, NY, USA, 427–434. https://doi.org/10.1145/2901790.2901890

[55] Yaxing Yao, Justin Reed Basdeo, Smirity Kaushik, and Yang Wang. 2019. Defending My Castle: A Co-Design Study of Privacy Mechanisms for Smart Homes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA. https://doi.org/10.1145/3290605.3300428

[56] Yaxing Yao, Justin Reed Basdeo, Oriana Rosata Mcdonough, and Yang Wang. 2019. Privacy Perceptions and Designs of Bystanders in Smart Homes. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–24.

[57] Eric Zeng, Shrirang Mare, and Franziska Roesner. 2017. End User Security & Privacy Concerns with Smart Homes. In *Proceedings of the Symposium on Usable Privacy and Security (SOUPS '17)*. USENIX Association, Berkeley, CA, USA, 65–80.

[58] Eric Zeng and Franziska Roesner. 2019. Understanding and Improving Security and Privacy in Multi-User Smart Homes: A Design Exploration and In-Home User Study. In *Proceedings of the USENIX Security Symposium (USENIX Security '19)*. USENIX Association, Berkeley, CA, USA, 159–176.

[59] Yu Zhai, Yan Liu, Minghao Yang, Feiyuan Long, and Johanna Virkki. 2014. A Survey Study of the Usefulness and Concerns About Smart Home Applications From the Human Perspective. *Open Journal of Social Sciences* 2, 11 (2014), 119. https://doi.org/10.4236/jss.2014.211017

[60] Serena Zheng, Noah Apthorpe, Marshini Chetty, and Nick Feamster. 2018. User Perceptions of Smart Home IoT Privacy. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 200. https://doi.org/10.1145/3274469

[61] Verena Zimmermann, Paul Gerber, Karola Marky, Leon Böck, and Florian Kirchbuchner. 2019. Assessing Users' Privacy and Security Concerns of Smart Home Technologies. *i-com* 18, 3 (2019), 197–216. https://doi.org/10.1515/icom-2019-0015

A CODING TABLE

In this section, we provide our coding tree based on the five categories of: 1) data sharing, 2) data collection, 3) information, 4) device providers, and 5) IoT devices. The coding tree does not include the codes that describe the data flow of individual IoT devices.

		Code	FE	KE	NE	Sum
Sharing	share	familiar owner	5	1	0	6
		familiar environment	6	2	0	8
		personal benefit	0	2	1	3
		interaction-based	1	0	2	3
		obvious data	2	3	3	8
		smalltalk	2	1	1	4
	limited	sensitive data	4	4	2	10
		unfamiliarity	1	3	1	5
	conditional	for high benefit	2	2	3	7
		for a benefit	15	10	10	35
		not even for a benefit	8	7	7	22
	no share	sensitive data	6	5	3	14
		unfamiliarity	0	1	2	3
		fear of consequences	0	2	2	4
		frequent visit	2	1	0	3
Collection	non-acceptance	data:consuming behavior-reason:disadvantages	1	2	1	4
		data:identity-reason:disadvantages	1	0	1	1
	acceptance	data:generic presence-reason:obvious	1	1	6	8
		data:preferences-reason:anonymous as visitor	1	0	0	1
Information		awareness	4	5	5	14
		possibility to consent	0	0	1	1
		personal interest	3	5	0	8
		purpose of collection	2	2	2	6
		data security / data protection	0	0	1	1
		transparency	5	7	7	19
		trust	2	2	0	4
		when data collection unexpected	0	3	2	5
		no information needed due to familiarity	3	1	0	4
Device Provider	no expectations	familiarity	1	0	0	1
		infrequent visit	0	1	0	1
	other	no trust in industry	0	0	2	2
		regulation by politics	0	1	0	1
	provider-related	device's capabilities	0	0	1	1
		controllability-self determination	1	1	3	5
		increase awareness	2	3	0	6
		security-related information	3	5	2	10
		transparency	2	4	8	14
		transparency-data practices	6	4	4	14