Defending My Castle: A Co-Design Study of Privacy Mechanisms for Smart Homes

Yaxing Yao
Social Computing Systems (SALT) Lab
School of Information Studies, Syracuse University
yyao08@syr.edu

Justin Reed Basdeo
Social Computing Systems (SALT) Lab
School of Information Studies, Syracuse University
jrbasdeo@syr.edu

Smirity Kaushik
Social Computing Systems (SALT) Lab
School of Information Studies, Syracuse University
smkaushi@syr.edu

Yang Wang
Social Computing Systems (SALT) Lab
School of Information Studies, Syracuse University
ywang@syr.edu

ABSTRACT

Home is a person's castle, a private and protected space. Internet-connected devices such as locks, cameras, and speakers might make a home "smarter" but also raise privacy issues because these devices may constantly and inconspicuously collect, infer or even share information about people in the home. To explore user-centered privacy designs for smart homes, we conducted a co-design study in which we worked closely with diverse groups of participants in creating new designs. This study helps fill the gap in the literature between studying users' privacy concerns and designing privacy tools only by experts. Our participants' privacy designs often relied on simple strategies, such as data localization, disconnection from the Internet, and a private mode. From these designs, we identified six key design factors: data transparency and control, security, safety, usability and user experience, system intelligence, and system modality. We discuss how these factors can guide design for smart home privacy.

CCS CONCEPTS

• **Security and privacy** → **Privacy protections**;

KEYWORDS

Smart Home, Internet of Things, Privacy, Co-Design

ACM Reference format:
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 CHI Conference on Human Factors in Computing Systems Proceedings, Glasgow, Scotland UK, May 4–9, 2019 (CHI 2019),* 12 pages.
https://doi.org/10.1145/3290605.3300428

1 INTRODUCTION

A smart home consists of different sensors, systems, and devices, which can be remotely controlled, accessed and monitored [7, 19]. The massive amount of data collected by Internet of Things (IoT) devices in a smart home allows entities to infer sensitive information without actually collecting them [9, 28]. Even seemingly innocuous data, such as home temperature and air conditioner status, could be used to determine whether a house is empty or not [23, 36]. In addition, people have expressed privacy concerns about smart homes, such as continuous data collection, sharing, and even misuse [6, 40, 42]. Privacy has thus been identified as a road blocker in the wide adoption of smart homes [18, 21].

To mitigate these concerns, different privacy mechanisms have been proposed, e.g., introducing noise to shape the smart home network traffic to prevent data inference [2]. However, little is known about what kinds of smart home

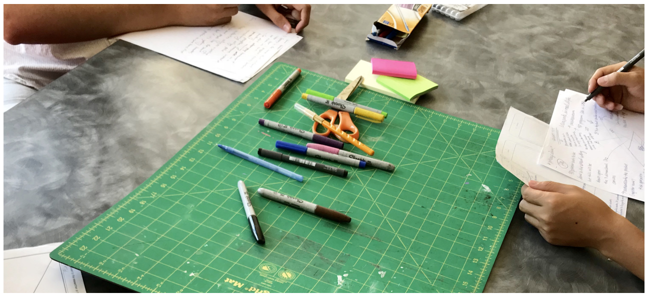

Figure 1: A photo was taken during one study session.

CHI 2019, May 4–9, 2019, Glasgow, Scotland UK

ACM ISBN 978-1-4503-5970-2/19/05…$15.00
https://doi.org/10.1145/3290605.3300428

privacy controls people desire. This is an important question to answer because privacy designs that address these desires are likely to be adopted.

To answer this question, we adopted a co-design approach to empower end users and engage them directly in the design process. Co-design [34] is a collaborative design approach in which stakeholders–such as researchers, designers, and users or potential users who are considered as "experts of their experiences" [38]–share their perspectives and cooperate creatively to generate new designs [35]. Kraemer and Ivan advocated that privacy issues in the smart home context should be approached by considering different stakeholders [20]. In our work, we collaborated closely with many groups of participants with diverse backgrounds in designing privacy mechanisms through a series of co-design sessions.

Our main contribution is that we identified six key design factors from our participants' designs of privacy mechanisms for smart homes. These factors include data transparency and control, security, safety, usability and user experience, system intelligence, and system modality. They reflected our participants' expectations in privacy mechanisms for smart homes and can be used as a good starting point to think about the design space of smart home privacy mechanisms.

2 RELATED WORK

Smart Home Privacy Concerns and Risks

Prior literature identified a number of privacy and security risks of smart homes. Arbo et al. pointed out the possibility of identity theft and device reconfiguration, suggesting the need for effective malware management [3]. With an experiment, Apthorpe et al. demonstrated how to infer sensitive user interaction with smart home devices through network traffic analyses with reasonable accuracy [1, 2]. A risk analysis of a smart home automation system by Jacobsson et al. pointed out that human-related risks (e.g. poor password selection) and software component risks (e.g., unauthorized modification of functions in the app) were the riskiest ones [17].

End users' privacy concerns have also been examined. By understanding people's mental model of how smart homes work, Zimmerman et al. uncovered participants' privacy concerns about hacker attacks and data abuse [45]. Brush et al. identified four barriers that defer the broad adoption of smart homes, such as "difficulty achieving security" in smart door locks and cameras [6]. Zeng et al. identified a number of concerns people have, such as continuous video recording, data collection and mining, network attacks on local networks, and account hacking [42]. However, people tend to outweigh cost and interoperability over privacy and security [42]. Worthy et al. found that the fewer trust participants had towards the entities who used their information, the greater control over information collection participants desired [40]. Zheng et al.'s study, on the contrary, found that their participants assumed their privacy is well protected because they trust their smart home device manufacturers [43].

Other studies focused on specific smart home devices. Malkin et al.'s survey about smart TVs revealed their respondents' uncertainty of data collection and usage as well as the common nonacceptance of data being re-purposed or shared with third parties [24]. McReynolds et al.'s study on smart toys unveiled parents' concerns about the toys' recording and data sharing abilities and children's concerns about being heard by their parents [26]. Lau et al.'s study about smart speakers found that users' rationales behind a lack of privacy concerns could lead them to serious privacy risks [22].

Smart Home Privacy Mechanisms

Researchers have proposed various solutions to mitigate privacy concerns and risks in smart homes. For instance, Apthorpe et al.'s solution decreased the inference of sensitive user activities by introducing a minimum amount of noise data to shape the smart home network traffic [2]. Datta et al. developed a Python library for IoT developers to easily implement privacy-preserving traffic shaping [11]. By injecting synthetic network packets, Yoshigoe et al.'s solution obscured the real network traffic and reduced potential privacy vulnerabilities [41]. Wang et al. built a live video analytic tool for denaturing video streams by blurring faces according to user-defined rules [39].

To reduce improper access to users' data, Moncrieff et al. developed a tool to dynamically manage access privileges based on a number of contextual factors in smart home surveillance (e.g., occupants' location and content of ongoing conversation) [29]. Arbo et al. proposed a framework to ensure data security for smart home devices by providing dynamically generated policies and interfaces in which end users could use to set up their privacy zones [3]. Chakravorty et al. designed a system to collect and store users' data, then only allow users to access their data upon successful re-identification [8].

To increase transparency and user control, Das et al. proposed an infrastructure for IoT devices and sensors to send personalized privacy notice and choice based on individual users' preferences [10]. McReynolds et al.'s study on smart toys suggested that toys should effectively communicate with both parents and children that toys could record [26].

More broadly, to ensure an overall safe environment of smart homes, Lin et al. suggested that auto-configuration support should be developed for smart home network so that whenever a new device is plugged into the network, the supporting system could auto-configure itself and find the most secure settings for the new devices, such as security protocols and essential firmware updates [23].

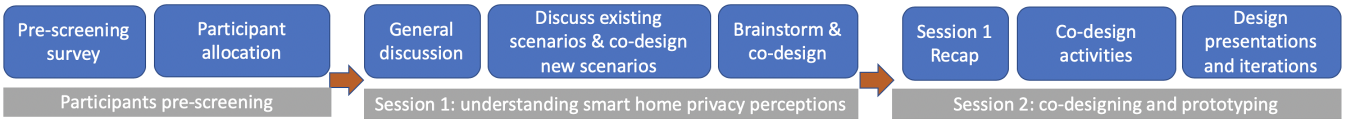
Pre-screening survey
Participant allocation
Participants pre-screening
General discussion
Discuss existing scenarios & co-design new scenarios
Brainstorm & co-design
Session 1: understanding smart home privacy perceptions
Session 1 Recap
Co-design activities
Design presentations and iterations
Session 2: co-designing and prototyping

Figure 2: The flow of our co-design study, including its various components.

The commonality of the above mechanisms is that they were proposed or developed solely by experts or researchers. Our work focuses on end users' needs and perspectives, helping fill the gap in the smart home literature between studying users' privacy concerns and designing privacy tools only by experts.

3 METHOD

To explore how people desire to protect their privacy in the context of smart homes, we conducted a set of co-design sessions with a total of 25 participants. Figure 2 shows our study flow including participant recruitment and two co-design sessions. Each session took about 1.5 hours and each participant was paid $15 for each session they participated in. Our study was approved by our university IRB.

Participants

Recruitment. We recruited our participants primarily through word-of-mouth, Craigslist, local community centers, libraries, and senior citizen centers. We framed our study as "a design study for smart home technology" and did not mention the word "privacy" to avoid any potential bias. We designed a pre-screening survey to get participants' demographic information and their experiences with smart home devices.

Participant data. The ages of our 25 participants ranged from 22 to 76 (mean: 41). 13 participants were cisgender female and the other 12 were cisgender male. They had various occupations, such as university staff, librarians, students, software engineers, retired workers, a security guard, a researcher, and a plumber. They were categorized into three types of participants based on their levels of experiences with smart homes: 12 participants owned smart home devices (*users*), 7 participants were interested in buying smart home devices (*interested users*), and 6 participants did not use or plan to buy smart home devices (*non-users*).

Study groups. Participants were divided into five groups (Group A, B, C, D, and E) primarily based on their schedules and levels of experiences with smart home devices. Each group had four to six participants (A:6, B:4, C:5, D:4, E:6). Each session had at least four participants except that Session 2 of Group D only had two participants due to schedule conflicts. Group E consisted of participants from a senior citizen center. Due to their mobility needs, we chose to conduct the study in their center with the participants from that center only. All other sessions were conducted in our lab.

All participants were invited to both sessions, however not everyone could attend both sessions due to practical constraints (e.g., conflicting schedules). To mitigate this issue, similar to [25], we started each Session 2 with a 15-minute recap of the discussion from the corresponding Session 1 (e.g., pros and cons and privacy concerns of smart home devices) to bring all participants to the same page. In the end, ten participants completed both sessions. Eleven participants only did Session 1 and four participants only did Session 2.

Session 1

The goal of the first session was to understand participants' privacy and security concerns of smart homes and to conduct the initial brainstorming and design. Each session started with a round-table introduction. We then asked each participant to talk about their experiences with smart home technologies and general perceptions. We then provided a working definition of a smart home, "a home that has different sensors, systems, and devices, which can be remotely controlled, accessed and monitored" based on the literature [7, 19]. We showed and explained pictures of a few smart home devices (e.g., voice assistants, smart thermostats, security cameras, and smart toys) to illustrate this smart home definition and potential uses of these devices [27].

Next, we asked our first group-based discussion question, "what are the pros and cons of these devices in your opinion?" This question was meant to frame the discussion in a neutral/balanced manner. In our pilot study, we found that our participants were overly excited about smart home, tended to fixate on the pros, and hardly considered the cons. To encourage participants to think more about the risks, we added a follow-up question in the actual study that asked, "have you experienced or heard of any negative incidents of smart home technologies, and what can potentially go wrong with smart home technologies?" This question did not prime our participants to only consider the negative aspects because we asked the pros first and they mentioned many pros. The careful consideration of both benefits and risks helped them to consider the trade-off in later co-design activities.

The next two activities were scenario-based because smart home devices can be used in various scenarios or for different purposes and privacy is highly contextual [31, 32]. As such, we hoped to provide opportunities for our participants to explore nuances of smart homes and their contextualized privacy implications. The first activity was a role play. We presented three scenarios adapted from the literature: (1) an Amazon Echo records a conversation between a couple and sends it to other people [37], (2) a security camera monitors a senior citizen's well-being at home in case of emergencies [44], and (3) a smart toy records and processes a child's conversation with it in order to respond, but also allows the parents to hear the conversation via a mobile app [26]. These scenarios were chosen to represent different devices, social relationships and power dynamics in the home (e.g., a couple and a co-worker, an older adult and an adult child, and young children and parents). In each scenario, we designed two to four roles for our participants to choose from. Once each participant picked a role, they discussed the potential privacy issues from the standpoint of the role.

The next activity was to co-create smart home usage scenarios. We encouraged our participants to work in groups of two or three. Each group chose a specific smart home device and co-created a usage scenario of that device with one or two researchers. We used six questions to guide the scenario creation process, i.e., what the device is, where the device is used, who uses the device, when to use the device, why uses the device, and how the device is used. Participants then presented the scenario and discuss any potential privacy implications in that scenario.

Through the above two activities, our participants discussed a wide range of usage scenarios and privacy issues of smart homes. We then moved on to the co-design activity. Specifically, we asked our participants to brainstorm their desired ways to mitigate these privacy issues and draw their design ideas. We provided a number of creation tools (e.g., colored papers, post-it notes, color pens). A student designer was also on site to provide help with sketching if needed. We deliberately asked our participants to work individually, think outside of the box, and consider different kinds of potential solutions. We also explained that the solutions could be futuristic and speculative without considering the status quo. Each participant then presented their ideas to the group.

Session 2

The goal of Session 2 was to continue the co-creation of privacy mechanisms, moving from ideation to creation of prototypes. We started the session by recapping the discussion from Session 1. To help our participants understand different forms of prototypes, we show various examples (e.g., paper prototypes of smartphone screens and a website). Similar to Session 1, we provided different creative tools and had a student designer on site to help them draw.

Each participant had about an hour to work on their prototypes. We encouraged them to discuss their ideas with other participants and the researchers, and then to create the design individually. Then each participant presented and discussed their prototypes with the group.

Data Analysis

Transcriptions and notes. All sessions were audio-recorded upon participants' permissions. The recordings were transcribed and then analyzed using a thematic analysis [5] by three co-authors. First, we immersed ourselves in the data by reading through the transcripts multiple times. Then we coded one transcription together at the sentence level to develop an initial codebook. Second, we independently coded the same transcription of another session using the codebook. We added new codes to the codebook in that process. Once finished, we compared and discussed our coding, and converged on an updated codebook. The inter-coder reliability was 0.91 (Cohen's Kappa), which is considered good [14]. Next, we coded the rest of the data using the updated codebook, which contains more than 100 unique codes, such as "self-driving car risks," "voice assistant authentication," "home context," "block data collection," "sharing decision," and "intrusion detection." Once finished coding, we grouped all codes into several themes, such as "data transparency and control," "security," "safety," "usability and user experience," "system intelligence," and "system modality." We examined and ensured the codes were assigned to the correct theme.

Image data. We collected participant drawings of their design ideas. Following Poole et al.'s methodology [33], three co-authors coded all elements in every design, including all components involved (e.g., stakeholders, devices, users), information flow, context, as well as other visual elements (e.g., icons, symbols, colors). Over 80 codes emerged from the analysis, and all the codes were grouped into the aforementioned themes in the analysis of audio transcriptions and notes. The inter-coder reliability was 0.84 (Cohen's Kappa).

4 RESULTS

Our work contributes to new understandings of how people conceptualize privacy control mechanisms for smart homes. Our participants started with creating their own smart home device usage scenarios. They created a wide variety of usage scenarios covering different devices and purposes (e.g., smart security cameras for home safety, smart doorbells and locks for remotely locking/unlocking doors, a smart fridge to automate food refill and alert food expiration, and a smart robot to support indoor navigation for people with visual impairments). This activity allowed our participants to explore and discuss possible ways of using smart home devices and potential privacy implications. These scenarios also served as a basis for the subsequent co-design of smart home privacy controls. Next, we will turn to our participants' smart home privacy designs, focusing on the major factors considered in their designs and identified via our thematic analysis. Table 1 shows an overview of these factors. It is worth noting that

Table 1: The six factors identified from our participants' designs of smart home privacy controls.

Factors	Examples
Data Transparency & Control	- Transparency and user awareness - Data localization - Disconnection from the Internet - Other user controls of data
Security	- Authentication of multiple users - Access control - Network intrusion detection
Safety	- Notification of physical break-in
Usability & UX	- Considerations of user characteristics - Considerations of user effort
System Intelligence	- Context detection - Personalization
System Modality	- Hardware devices - Apps, modes, policies

these factors are not mutually exclusive. One design might consider multiple factors. We present these factors below.

Data Transparency and Control

A major privacy concern shared by our participants was smart home devices collecting data about them. They created various designs to increase the transparency of data collections and user control over their data. Seventeen participants considered this factor (P2-6, 9-10, 12, 15-21, 24-25).

Transparency and user awareness. Seven participants' designs (P2-3, 6, 12, 15, 18-19) were centered around improving transparency of data collection and usage of smart home devices. For example, P15 designed a transparency feature for a self-driving car. She considered the car part of the smart home because the car is often parked/charged at home and she can control the car (e.g., start the engine) remotely using voice assistants (e.g., Google Home). However, she was concerned that the car manufacturer might collect her car usage data (e.g., when she used the car, where she had been to) and then use that data to predict her future activities. To address these concerns, she designed the car with two modes, an invisible mode and a visible mode. When she wishes not to be tracked, she can turn on the *invisible* mode (e.g., by plugging in a dedicated USB drive to the car) to hide her activities. In contrast, under the *visible* mode (default mode), her driving data can be tracked but she can use an app interface to see what data about her has been collected.

P15 explained, *"so the visible basically tells you what you have done with this car, like a transparency tool... [the tool] can also make sense of how the manufacturer uses the data. Like they can infer whether I'm a night person or not to increase my insurance payment."* (P15) Her design of the visible mode provides more transparency about the car's data collection and usage practices. However, she suggested this feature should be provided by third-party companies because she felt the car manufacturers might not tell the truth.

It is also worth noting that, our participants considered the purposes of data collection in their smart home scenarios when designing the privacy controls. For instance, even in the invisible mode of P15's design, she would share when/where she uses the car with a trusted third party (e.g., for better navigation purpose) but would not with the car manufacturer. In P3's design, the security camera monitors his home for safety (i.e., purpose), but the associated app does not show the actual images or videos and only offers text descriptions thereof to mitigate privacy risks.

Data localization. A common element across seven participants' designs (P2, 4-5, 15-16, 20, 24) was data localization, the idea that smart home devices store and process the collected data locally as opposed to sending the data to a remote server. For example, P16 designed a smart door lock with a fingerprint reader to improve the safety of her home. Since the fingerprint reader collects her biometric information, she designed an additional privacy feature to protect her fingerprint data by only storing it locally in the lock. She explained, *"the fingerprint will be stored onsite only. There is no need to have it connected to anything. If you didn't have your original key and you didn't get in with your fingerprint for some reason, the only thing that the company can do is complete pledge it and you would start it as a brand-new device because they would not have access to get into that."* (P16) According to P16's design, her fingerprint data will reside in the lock, but the smart lock is still network-connected because the company could remotely reset the lock if the user lost her key or the fingerprint reader stopped working. However, we note that remotely resetting locks can pose security risks.

Disconnection from the Internet. This privacy mechanism means disconnecting smart home devices from the Internet, essentially working in an offline manner. Five participants' designs (P2-4, 24-25) included this idea. For instance, P2 was concerned about home security cameras collecting personal data and storing these data in cloud storage, which may not be secure. To address these concerns, he proposed a design of a physical lock that could be plugged into a security camera to protect his personal data. He explained, *"I think what people need is something like a lock that can be plugged into the security camera to lock our data like gender or activities. Now they [security cameras] are using cloud services like the iCloud to store my personal data, but I don't know whether they are secure or not because they are stored at some other* *place, so if I have my own device without the Internet, that is safer. It's like a physical control and my things are stored only in my place."* (P2) Several interesting ideas are behind this design. First, the lock is intelligent in selectively filtering out certain types of data (e.g., gender). Second, the lock can disconnect the security camera from the Internet/network. A defining characteristic of smart home devices or IoTs more broadly is their Internet connectedness, which often supports data transfer, remote control, and other system intelligence (e.g., predictions). However, here we see P2's desire to directly control (enable/disable) the Internet connectedness. By disconnecting the security camera from the Internet, P2 felt that he has more (at least perceived) control over the data. Third, the lock (as a standalone hardware device) is physical, which affords more tangible control.

Other user controls of data. Besides data localization and disconnection from the Internet, nine participants (P5, 9-10, 12, 15, 18, 21, 24-25) desired explicit controls of their data, from preventing data collection to deleting collected data. The aforementioned P15's example of the invisible mode of a self-driving car was designed to prevent car manufacturers from collecting data about her car activities. In comparison, P5 designed a conceptual model of smart home privacy mechanisms and emphasized that a key aspect of his model was users' ability to delete data. He explained, *"the user should have a hardware option to delete data. So they don't have to necessarily go to the software to delete it."* (P5) He believed that users should be able to delete the data collected about them and the deletion feature should be implemented as a hardware option (e.g., a physical button on the smart home device), which would be easier to use than a software control.

Security

Another underlying factor of participants' designs was related to security, including aspects such as authentication of multiple users, access control of user data, and network intrusion detection. Twenty participants (all participants except P15, 18-19, 24-25) considered this factor in their designs.

Authentication of multiple users. Eleven participants (P1, 4-5, 10-13, 17, 21-23) spoke to the social relationships and power dynamics in homes where there could be multiple users sharing one device. They emphasized the importance of enabling proper authentication in order to protect each family member's privacy. For example, P13 was concerned that other members in the household might be able to access her credit card information and order food from the smart fridge. To address this concern, she incorporated voice recognition in her design as an authentication mechanism for the smart fridge. She explained, *"even if someone hacks your details about the credit card to make payment, but it will still need your voice to recognize and authenticate that transaction. So that can't happen unless you do it yourself. Even if someone has credit card details, the transaction won't go through."* (P13) While P13's voice could uniquely identify/authenticate her, she did not speak to the possibility where someone else might record and replay her voice to impersonate her (i.e., replay attacks).

Access control. In addition to authentication, authorization or access control of who can access what data was another security feature that 16 participants (P1-11, 14, 16-17, 20, 23) considered in their designs. For instance, security cameras (e.g., Nest Cam [30]) often allow anyone who logs into the compatible mobile app to see the same video content. However, P3 wanted to give users different access rights, as shown in the left screen in Figure 3. He designed two modes. In the *online* mode, the app shows the video feed from the camera. In the *offline* mode, the app provides a textual description of the video feed (e.g., a person is walking), as shown in the middle screen in Figure 3. The access control (the right screen in Figure 3) determines who gets to use which mode. He elaborated, *"you can decide who should be in which mode from this access management page, so some will see the text description, some will see the live video."* (P3) This feature is similar to sharing location data at varying granularity (e.g., actual address vs. city) with different entities.

In addition, some participants also designed location-based access controls. For example, P1 explained that in his design of a home automation system, the app to control his smart home devices should have a *local* mode and a *remote* mode. He should have full access to all the functions and data only when he is physically at home, which triggers the local mode. In comparison, when he is away from home, the system will enter the remote mode and should only give him partial access to the devices and none of his data should be transmitted through the Internet. We are not aware of any existing home automation products that support this feature.

Network intrusion detection. Another security feature included in three participants' designs (P6, 22-23) was the ability to detect external intrusions into the smart home network. For example, P6 had a technical background, and he

Online mode
offline mode
✓Fast Access
Security concern
✓Auto Alarm
✓Slow Access
SELECT
SELECT
private mode ∞
(text mode)
2018-07-17-13:00
: No action detected
C. Ronaldo
2018-07-01 21:01
L. Messi
K. Mbappe
T. Henry

Figure 3: The online and offline modes of security cameras (P3)

was particularly concerned that hackers might hack into his home network and steal his information, that his neighbors could connect to his home network and invade his privacy, or his devices could send his personal information to third parties other than the device manufacturers. To address these concerns, he designed a smart router with a built-in firewall and an app that worked with the smart router, which could be used to notify users whenever an outside intrusion was detected. He elaborated, *"the app will be able to track data sending from each device and its destination. If there is anyone who is trying to penetrate your network or trying to use your data, collect your data, this app should send you identification. You can probably reboot the device from the app to stop, kind of like a filter."* (P6) This smart firewall would be able to track each smart home device's data flow and notify users of any third party attempting to collect data.

Safety

Since our participants designed for the home environment, safety was also a concern. Twelve participants (P1, 7-8, 11, 15, 17-18, 21-25) considered this factor. For example, our participants expected that security cameras or voice assistants should be able to notify either the homeowner or the police department if someone broke into their house. If the users were in an emergency and needed help, they should be able to call for help quickly from these devices. These features are often already supported by existing products. Some participants also tried to ensure the physical safety of the home while preserving people's privacy. For instance, P18 was concerned about the unknowingly recording of passersby by doorbell cameras. She then created a long list of key points that should be written into policies. She explained, *"I suggested that it could be first a law by the government that owners have to somehow make other people somehow aware whether that's a sign that says you're being recorded."* (P18) While the doorbell cameras can arguably help improve people's (safety) awareness of their home door area, P18 was concerned about the privacy of other people (e.g., passersby). By calling for legislation that requires a clear notice of such recording practice, P18 attempted to strike a good balance between homeowner safety and passersby privacy.

Usability and User Experiences

Many participants explicitly considered the usability of their privacy designs, ensuring that users have good user experiences with the designs. Twelve participants (P3, 11-15, 17-19, 21-23) considered this factor in their design. There were two broad categories of usability considerations: user characteristics and user effort.

Considerations of user characteristics. Seven participants (P11, 13, 18-19, 21-23) took people's characteristics (e.g., abilities) into account when designing their privacy mechanisms, hoping to make their designs more inclusive to a wide range of users. For instance, P11 designed an in-home robot, which could help people with various tasks in their homes. He then designed hardware access control interfaces to manage who could access the robot remotely. In the group discussion, P13 asked the following questions about P11's design, *"if the person is blind, or is it like an elderly person who cannot walk?"* (P13) These questions prompted P11 to reconsider his design. P11 then reduced the number of buttons in the interface so that it might be easier for a variety of users (e.g., children, older adults). Similarly, P18 originally designed an authentication mechanism for Amazon Echo using a physical fingerprint reader. Later he added a voice recognition mechanism for authentication because he realized that a physical fingerprint reader may be impossible or hard to use for people who lost their fingers or who have mobility impairments.

Considerations of user effort. Another usability consideration was the amount of effort required from users to utilize the privacy designs. The majority of design ideas were based on automation (e.g., users receiving automatic alerts about their information being used). However, eleven participants' designs (P3, 11-15, 17, 19, 21-23) intentionally required explicit user effort. For example, P19 designed an improved privacy policy (summaries of most important points) for a smart thermostat and he believed that companies should be required to show the policy and users should be required to read these policies to understand the data collection. However, we note that people tend not to read privacy policies.

System Intelligence

Twelve participants (P1, 4-9, 13, 15, 21-23) considered system intelligence in their design. Among participants' privacy designs, we noticed two types of system intelligence: context detection and personalization.

Context detection. Since homes can have various social relationships, contexts, and thus privacy implications, six participants' designs (P1, 9, 13, 21-23) included a component of automatic context detection. For instance, P9 designed smart toys for her children but was concerned that her sensitive data might be accidentally recorded by these toys. For instance, she might be calling the bank with her credit card information while her children play with the toys, which may record and leak her private information outside the house. To address this concern, she embedded a context detection feature as part of her design. She explained, *"like when we want to have a private discussion, they [smart toys] are not allowed to [record]."* (P9) She expected the toys to be able to automatically detect when she is having a private conversation and the toys will pause their recording. Similarly, P23 designed a security camera that can automatically detect that she is not at home and start recording in the home.

Personalization. Since the home might have multiple people with different needs, twelve privacy designs (P1, 4-9, 13, 15, 21-23) were personalized. For instance, P21, a senior citizen who lived with a portable oxygen concentrator, expected that her daughter can access her security camera to check on her well-being. However, P21 desired personalized preferences in terms of when her daughter can access the camera feed and when she cannot. P22 echoed her support, *"if I don't want them to see certain things, I can deny them. Don't have them on camera when I do this, this and this. I got company, I'm eating ice cream, don't bother, that'd be good. Program it so that we don't have to worry about it. Like an alarm, you can set an alarm based on what you are doing. That's what a true friend would do. Let Alexa be your true friend. Tweak it up!"* P22 expressed her desire of setting her personalized preferences of access control via a voice assistant, which then can automatically enforce these preferences.

System Modality

We observed four forms or modalities of how participants' privacy designs were embodied: hardware devices, apps, system modes, and policies. These modalities are not mutually exclusive. Some privacy designs had two or more modalities.

Hardware devices. Ten participants' designs (P2, 6-9, 11-12, 14-16) were proposed as hardware devices, such as P2's design of a physical lock for security cameras, P6's design of a smart router, and P15's design of a USB device for self-driving cars. In some cases, our participants intentionally designed their privacy mechanisms as a hardware solution. For example, P15 explained, *"it is just a USB, you plug it in, it will record the data, you can plug it into a computer to read...it's small, I can take it with me and plug it in whenever I want to hide my activities...I don't know whether it is possible to connect just using Bluetooth. So with the Internet, it can upload the data by itself, but with the Bluetooth, it can only transmit data from the device to your phone, then your phone can analyze the data itself, so the USB is safer."* P15's choice of a USB was due to its portability and its perceived security (no connection to the Internet).

Apps. Another common modality was a mobile app, often features of the mobile app associated with the smart home device. Twelve participants' designs (P1, 3-5, 9, 12, 14-15, 17, 21-23) took the form of an app. We have presented examples, such as P3's privacy design of the security camera app, and P15's privacy design of the app for self-driving cars.

Modes. Four privacy designs (P3-4, 12, 15) were envisioned as system modes in hardware devices or mobile apps, such as P3's design of online and offline modes for security camera and P15's design of visible and invisible modes for self-driving cars. These modes were often binary, privacy mode vs. regular mode. They mapped to some participants' coarse categorizations of privacy implications (e.g., I need privacy in this case). Some of them explicitly mentioned the incognito mode (of the Google Chrome browser), which likely inspired their designs.

Policy. In addition to technological solutions, six participants' privacy designs (P10, 18-20, 24-25) were in the form of laws and/or policies, for instance, P18's example of legislation, which would require smart doorbell cameras to clearly notify passersby that the cameras can record them.

5 DISCUSSION

Smart Home Privacy

The home context is complicated for privacy. First, smart home privacy covers not only information privacy (e.g., data collection and sharing) but also physical privacy (e.g., the privacy of the physical space of homes). Our participants paid attention to both types of privacy in their privacy designs (e.g., data transparency, safety).

Second, the complex social relationships and power dynamics in a home, such as parents and children, brothers and sisters, husband and wife, owners and guests, patients and remote doctors [15], can significantly affect whose privacy is at risk or how privacy can be enacted. Many privacy designs in our study supported multiple user accounts which have been explored for shared home computers [13], but also included multi-user authentication and access control.

Third, different social relationships may suggest varying privacy norms [31]. For example, having visitors in a home changes the social context of the home and its privacy norms. Homeowners might choose not to say things in front of their visitors. Similarly, if the smart home devices record or process the conversations in the home, visitors may feel their privacy is violated. An open question for designing smart home privacy mechanisms is, *whose privacy should be protected and who should make the decision?* While most of the participants' designs were for people who live in the home, we saw some cases where the privacy of other people (e.g., passersby) was considered.

Design Implications

Next, we will discuss how the list of design factors we identified from our participants' privacy designs can be used to guide the design of smart home privacy mechanisms.

Data transparency and control are relevant whenever smart home devices collect data and/or can infer data about people in the home and around the home (e.g., passersby). While notice and choice are well-respected privacy principles, how to best provide and implement them is still an open question for smart homes. In terms of notice, our participants desired more transparency about what data individual smart home devices, as well as the smart home system as a whole, can collect, infer, share and use about them. Therefore, privacy designs should be considered at both the individual device level and the whole system level (e.g., P6's design of a smart router that monitors the entire smart home system).

In terms of user control, many participants desired data localization, the ability to have the smart home devices store and process the collected data in the devices locally. While client-side data storage/processing has been proposed as a privacy-enhancing technique (e.g., in targeted advertising [4] and recommender systems [16]), most smart home devices rely on servers and cloud services to store and process the collected data [7]. In fact, some proposed mechanisms in the literature also require cloud storage [12], which conflicts with our participants' desires. We suggest that designers should consider data localization as a possibility.

In addition, many participants incorporated the idea of disconnection from the Internet in their privacy designs because they felt it will give them a peace of mind because their data cannot leak out via the Internet. We note that this is concerned with the public Internet rather than the private home network (Intranet). This idea challenges a typical assumption that all smart home devices are Internet connected. Do these devices always need to connect to the Internet and should they pause their data collection and sharing if users demand so? We believe that these are important questions that designers should consider. We also note that just because devices can disconnect from the Internet does not mean they cannot collect and send data after they resume Internet connections. Fundamentally, this idea is about giving users the option to say no to data collection and sharing. Disconnection from the Internet is a simple concept that people can understand and perceive better privacy/security.

Security is closely related to privacy. Our participants considered different kinds of security attack scenarios, ranging from other members of the home accessing their data to hackers breaking into the home network (gaining control of their devices and/or stealing their data) to the devices sending their data to external third parties. In response, our participants' designs covered multiple user authentication, authorization (access control of who can see what data), and network intrusion detection. Our participants were particularly concerned about information related to their health, finance, gender, location, and activities. Most of the current smart home devices lack these security features. We recommend designers to consider these options to address users' security concerns, e.g., if the devices allow direct interactions with users, then user authentication and authorization should be considered.

Safety was a natural concern for the home context. Many participants' designs included safety features (e.g., security cameras identifying suspicious activities). We recommend designers to consider these safety features, but more importantly, we encourage designers to think about whether safety and privacy might be in conflict. For instance, in P18's example of a doorbell camera, homeowner safety and passersby privacy might be at odds. How to reconcile when these two values conflict is another open question for further research.

Usability and user experiences are arguably important for any user-facing design. Our participants desired simple and easy-to-understand privacy mechanisms, for instance, the feature of disconnection from the Internet. In addition, they paid attention to the diversity of users and their varying needs as well as the amount of user effort required to use the designs. Our suggestion here is that designers should consider how to make their design more inclusive to various user groups and how to reduce user effort to use the privacy controls (e.g., designing privacy-friendly default settings).

System intelligence in our study covers automatic context detection and user personalization. While context-aware computing has been extensively studied, some of the designs included intelligent context detection that is currently hard to implement (e.g., security cameras automatically detecting and describing what is happening in a home). Supporting users' personalized privacy preferences has been explored in the IoT space (e.g., [10]) and designers should consider supporting this feature in their smart home privacy designs.

Lastly, *system modality* presents the form(s) in which these privacy designs are embodied. Our participants covered four modalities: hardware devices, apps, system modes, and policies. Designers should consider this question of modality because it could influence other aspects of their privacy designs, for instance, usability and user experiences. Some of our participants' privacy designs were deliberately envisioned as hardware controls (e.g., a USB device to turn on the invisible mode) due to its perceived ease of use and portability. Many designs were also based on binary modes (visible/invisible, or online/offline). This binary model is easy to understand and use in part because people have experiences with similar models in other domains (the private mode in web browsers).

Policy Implications

Some participants designed privacy policies (e.g., P18's suggested policy on smart doorbell cameras). They believed that the government should play an important role in ensuring device manufacturers behave appropriately, for instance, what they are allowed and not allowed to do. Our participants also discussed the following scenario: if a user encountered some negative incidents (e.g., robbery) due to data collection or sharing by the device manufacturers (e.g., the user's personal information was collected and leaked to the wrong hands, then the user's daily schedule was inferred), will the manufacturers be held accountable? To what extent current privacy laws such as the European Union (EU) General Data Protection Regulation (GDPR) and the California Consumer Privacy Act address these questions remains to be seen.

Reflections on Participants' Privacy Designs

Our co-design study aimed to give voice to people (users and non-users of smart homes), who are often not included in the design of privacy controls. Our participants contributed many novel ideas, such as stand-alone hardware devices as privacy controls (e.g., a physical lock for security cameras), seamless identification and authentication of multiple users, and automatic context-based personalized privacy controls (e.g., a smart toy selectively pausing its recording based on the detected context and a user's contextual preferences).

However, our participants' privacy designs also have limitations. First, since most of our participants were not technically savvy, so their designs did not cover all the privacy-enhancing techniques found in the literature, for instance, adding noise to the home network traffic to reduce data inferences [2]. Second, their designs did not address all potential privacy risks in smart homes, for instance, the risk of secondary use of data (using the collected data for a different purpose) [24]. Their designs also did not address the case that the manufacturers collected users' data for a reasonable purpose but then shared the data with third parties. Third, some of their designs are currently hard to implement (e.g., security cameras providing real-time textual descriptions of the video feed). However, this was by design because we did not want to limit our participants' creativity. Fourth, many designs could potentially pose privacy or security risks themselves. For example, the smart router monitoring the entire home network could be privacy intrusive itself. Remotely resetting a door lock could also have security risks.

All of these novel design ideas and concrete limitations suggest that when designing privacy mechanisms for smart homes, inputs from both users and privacy experts are needed.

Limitations of Our Research

Our study also has several limitations. First, we asked the participants "what can potentially go wrong with smart home devices," which may prime our participants to focus on the negative aspects. However, we believe that our participants were unlikely to make up issues because (1) participants were not required to answer this or any question, (2) the privacy concerns were often raised by multiple participants voluntarily, and (3) participants discussed similar concerns in other domains (e.g., web tracking). In addition, as we discussed in the method section, we asked about the pros and cons of smart homes first. Our results showed that our participants covered both the pros and cons in their considerations and their designs often reflected the trade-off between benefits and risks. Second, while our participants had very diverse backgrounds, we did not include anyone younger than 18. While our study did not focus on smart toys, some participants designed for smart toys. Having children as part of the co-design team would have been valuable for the privacy designs for smart toys. Third, all of our participants' designs were low-fidelity paper prototypes rather than interactive high-fidelity prototypes. Therefore, they might have missed potential challenges of their designs or opportunities to improve the designs. Fourth, our study focused on users and might have missed perspectives (factors) from other stakeholders such as device manufacturers. Our student designer who helped our participants with their design was not experienced in hardware designs, so the help was also limited.

Future Directions

The aforementioned limitations point to a few directions for future research. The limitations of our participants' privacy designs suggest that future work should not only continue to explore user-generated designs but also critically evaluate these designs in terms of their feasibility, usability, privacy, and security. These evaluations can shed light on how to effectively combine end users' ideas and expectations with experts' knowledge. These insights can then inform how these designs should be adapted and implemented in practice. Furthermore, future co-designs could consider educating users on privacy/security risks and countermeasures before starting the actual design. This might lead to additional designs. Lastly, we did not observe any differences in terms of perceptions or design factors among the users, non-users, and interested users. This might be due to the small sizes of different user groups in our study. Future work can further explore potential differences among various types of users.

6 CONCLUSION

Smart home devices are gaining momentum albeit with serious privacy challenges. We conducted a co-design study to understand how people desire to protect their privacy in the smart home context. From participants' designs of smart home privacy mechanisms, we identified six important design factors they considered: data transparency and control, security, safety, usability and user experience, system intelligence, and system modality. We discuss how these factors can guide the design of smart home privacy mechanisms. Future research should try to involve more stakeholders (e.g., device manufacturers) in the privacy design process, and further explore and evaluate user-generated privacy designs.

7 ACKNOWLEDGEMENT

We thank our participants for their valuable input. We thank Elias Greendorfer, Brianna Dym, Bryan Semaan, anonymous reviewers and shepherd for their thoughtful feedback. This work was supported in part by NSF Grant CNS-1464347.

REFERENCES

[1] Noah Apthorpe, Dillon Reisman, and Nick Feamster. 2017. A smart home is no castle: Privacy vulnerabilities of encrypted IoT traffic. *arXiv preprint arXiv:1705.06805* (2017).

[2] Noah Apthorpe, Dillon Reisman, Srikanth Sundaresan, Arvind Narayanan, and Nick Feamster. 2017. Spying on the smart home: Privacy attacks and defenses on encrypted iot traffic. *arXiv preprint arXiv:1708.05044* (2017).

[3] Abdullahi Arabo, Ian Brown, and Fadi El-Moussa. 2012. Privacy in the age of mobility and smart devices in smart homes. In *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom)*. IEEE, 819–826.

[4] Mikhail Bilenko, Matthew Richardson, and Janice Tsai. 2011. Targeted, not tracked: Client-side solutions for privacy-friendly behavioral advertising. In *Privacy Enhancing Technologies Symposium (PETS 2011)*.

[5] Richard E Boyatzis. 1998. *Transforming qualitative information: Thematic analysis and code development.* sage.

[6] AJ Brush, Bongshin Lee, Ratul Mahajan, Sharad Agarwal, Stefan Saroiu, and Colin Dixon. 2011. Home automation in the wild: challenges and opportunities. In *proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2115–2124.

[7] Joseph Bugeja, Andreas Jacobsson, and Paul Davidsson. 2016. On privacy and security challenges in smart connected homes. In *Intelligence and Security Informatics Conference (EISIC), 2016 European*. IEEE, 172–175.

[8] Antorweep Chakravorty, Tomasz Wlodarczyk, and Chunming Rong. 2013. Privacy preserving data analytics for smart homes. In *Security and Privacy Workshops (SPW), 2013 IEEE*. IEEE, 23–27.

[9] Federal Trade Commission et al. 2015. Internet of Things: Privacy & security in a connected world. *Washington, DC: Federal Trade Commission* (2015).

[10] Anupam Das, Martin Degeling, Daniel Smullen, and Norman Sadeh. 2018. Personalized Privacy Assistants for the Internet of Things. (2018).

[11] Trisha Datta, Noah Apthorpe, and Nick Feamster. 2018. A Developer-Friendly Library for Smart Home IoT Privacy-Preserving Traffic Obfuscation. In *Proceedings of the 2018 Workshop on IoT Security and Privacy*. ACM, 43–48.

[12] Ali Dorri, Salil S Kanhere, Raja Jurdak, and Praveen Gauravaram. 2017. Blockchain for IoT security and privacy: The case study of a smart home. In *Pervasive Computing and Communications Workshops (PerCom Workshops), 2017 IEEE International Conference on*. IEEE, 618–623.

[13] Serge Egelman, AJ Brush, and Kori M Inkpen. 2008. Family accounts: a new paradigm for user accounts within the home environment. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. ACM, 669–678.

[14] Joseph L Fleiss, Bruce Levin, and Myunghee Cho Paik. 2013. *Statistical methods for rates and proportions.* John Wiley & Sons.

[15] David Frohlich and Robert Kraut. 2003. The social context of home computing. In *Inside the smart home*. Springer, 127–162.

[16] Simon Gerber, Michael Fry, Judy Kay, Bob Kummerfeld, Glen Pink, and Rainer Wasinger. 2010. PersonisJ: Mobile, Client-Side User Modelling. In *International Conference on User Modeling, Adaptation, and Personalization (Lecture Notes in Computer Science)*, Vol. 6075. Springer Berlin / Heidelberg, 111–122. http://dx.doi.org/10.1007/978-3-642-13470-8_12

[17] Andreas Jacobsson, Martin Boldt, and Bengt Carlsson. 2016. A risk analysis of a smart home automation system. *Future Generation Computer Systems* 56 (2016), 719–733.

[18] Andreas Jacobsson and Paul Davidsson. 2015. Towards a model of privacy and security for smart homes. In *Internet of Things (WF-IoT), 2015 IEEE 2nd World Forum on*. IEEE, 727–732.

[19] Li Jiang, Da-You Liu, and Bo Yang. 2004. Smart home research. In *Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on*, Vol. 2. IEEE, 659–663.

[20] Martin J Kraemer and Ivan Flechais. 2018. Researching privacy in smart homes: A roadmap of future directions and research methods. (2018).

[21] Sathish Alampalayam Kumar, Tyler Vealey, and Harshit Srivastava. 2016. Security in internet of things: Challenges, solutions and future directions. In *System Sciences (HICSS), 2016 49th Hawaii International Conference on*. IEEE, 5772–5781.

[22] Josephine Lau, Benjamin Zimmerman, and Florian Schaub. 2018. Alexa, are you listening? privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. In *Pro. ACM Human-Computer Interaction CSCW*. ACM.

[23] Huichen Lin and Neil W Bergmann. 2016. IoT privacy and security challenges for smart home environments. *Information* 7, 3 (2016), 44.

[24] Nathan Malkin, Julia Bernd, Maritza Johnson, and Serge Egelman. 2018. "What Can't Data Be Used For?" Privacy Expectations about Smart TVs in the US. *European Workshop on Usable Security (EuroUSEC)* (2018).

[25] Roisin McNaney, John Vines, Jamie Mercer, Leon Mexter, Daniel Welsh, and Tony Young. 2017. DemYouth: Co-Designing and Enacting Tools to Support Young People's Engagement with People with Dementia. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 1313–1325.

[26] 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*. ACM, 5197–5207.

[27] Tiago DP Mendes, Radu Godina, Eduardo MG Rodrigues, João CO Matias, and João PS Catalão. 2015. Smart home communication technologies and applications: Wireless protocol assessment for home area network resources. *Energies* 8, 7 (2015), 7279–7311.

[28] Andrew Meola. 2016. How the Internet of Things will affect security & privacy.

[29] Simon Moncrieff, Svetha Venkatesh, and Geoff West. 2007. Dynamic privacy in a smart house environment. In *Multimedia and Expo, 2007 IEEE International Conference on*. IEEE, 2034–2037.

[30] Nest. 2018. Nest Cam. https://nest.com/cameras/nest-cam-indoor/overview/

[31] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Washington law review* 79, 1 (2004), 119–158.

[32] Helen Nissenbaum. 2009. *Privacy in context: Technology, policy, and the integrity of social life.* Stanford University Press.

[33] Erika Shehan Poole, Christopher A Le Dantec, James R Eagan, and W Keith Edwards. 2008. Reflecting on the invisible: understanding end-user perceptions of ubiquitous computing. In *Proceedings of the 10th international conference on Ubiquitous computing*. ACM, 192–201.

[34] Elizabeth B-N Sanders and Pieter Jan Stappers. 2008. Co-creation and the new landscapes of design. *Co-design* 4, 1 (2008), 5–18.

[35] Marc Steen, Menno Manschot, and Nicole De Koning. 2011. Benefits of co-design in service design projects. *International Journal of Design* 5, 2 (2011).

[36] Biljana L Risteska Stojkoska and Kire V Trivodaliev. 2017. A review of Internet of Things for smart home: Challenges and solutions. *Journal of Cleaner Production* 140 (2017), 1454–1464.

[37] The New York Times. 2018. Is Alexa Listening? https://www.nytimes.com/2018/05/25/business/amazon-alexa-conversation-shared-echo.html

[38] Froukje Sleeswijk Visser, Pieter Jan Stappers, Remko Van der Lugt, and Elizabeth BN Sanders. 2005. Contextmapping: experiences from

practice. *CoDesign* 1, 2 (2005), 119–149.

[39] Junjue Wang, Brandon Amos, Anupam Das, Padmanabhan Pillai, Norman Sadeh, and Mahadev Satyanarayanan. 2018. Enabling Live Video Analytics with a Scalable and Privacy-Aware Framework. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 14, 3s (2018), 64.

[40] Peter Worthy, Ben Matthews, and Stephen Viller. 2016. Trust me: doubts and concerns living with the Internet of Things. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. ACM, 427–434.

[41] Kenji Yoshigoe, Wei Dai, Melissa Abramson, and Alexander Jacobs. 2015. Overcoming invasion of privacy in smart home environment with synthetic packet injection. In *TRON Symposium (TRONSHOW), 2015*. IEEE, 1–7.

[42] Eric Zeng, Shrirang Mare, and Franziska Roesner. 2017. End User Security & Privacy Concerns with Smart Homes. In *Symposium on Usable Privacy and Security (SOUPS)*.

[43] Serena Zheng, Marshini Chetty, and Nick Feamster. 2018. User Perceptions of Privacy in Smart Homes. *arXiv preprint arXiv:1802.08182* (2018).

[44] Martina Ziefle, Carsten Rocker, and Andreas Holzinger. 2011. Medical technology in smart homes: exploring the user's perspective on privacy, intimacy and trust. In *Computer Software and Applications Conference Workshops (COMPSACW), 2011 IEEE 35th Annual*. IEEE, 410–415.

[45] Verena Zimmermann, Merve Bennighof, Miriam Edel, Oliver Hofmann, Judith Jung, and Melina von Wick. 2018. 'Home, Smart Home'–Exploring End Users' Mental Models of Smart Homes. *Mensch und Computer 2018-Workshopband* (2018).