More than Smart Speakers: Security and Privacy Perceptions of Smart Home Personal Assistants

Noura Abdi, ***King's College London;*** **Kopo M. Ramokapane,** ***University of Bristol;***
Jose M. Such, ***King's College London***

https://www.usenix.org/conference/soups2019/presentation/abdi

This paper is included in the Proceedings of the Fifteenth Symposium on Usable Privacy and Security.

August 12–13, 2019 • Santa Clara, CA, USA

ISBN 978-1-939133-05-2

Open access to the Proceedings of the Fifteenth Symposium on Usable Privacy and Security is sponsored by USENIX.

More than Smart Speakers: Security and Privacy Perceptions of Smart Home Personal Assistants

Noura Abdi
Department of Informatics
King's College London
United Kingdom
noura.abdi@kcl.ac.uk

Kopo M. Ramokapane
Bristol Cyber Security Group
University of Bristol
United Kingdom
marvin.ramokapane@bristol.ac.uk

Jose M. Such
Department of Informatics
King's College London
United Kingdom
jose.such@kcl.ac.uk

Abstract

Smart Home Personal Assistants (SPA) such as Amazon Echo/Alexa and Google Home/Assistant have made our daily routines much more convenient, allowing us to complete tasks quickly and efficiently using natural language. It is believed that around 10% of consumers around the world already own an SPA, and predictions are that ownership will keep rising. It is therefore paramount to make SPA secure and privacy-preserving. Despite the growing research on SPA security and privacy, little is known about users' security and privacy perceptions concerning SPA complex ecosystem, which involves several elements and stakeholders. To explore this, we considered the main four use case scenarios with distinctive architectural elements and stakeholders involved: using built-in skills, third-party skills, managing other smart devices, and shopping, through semi-structured interviews with SPA users. Using a grounded theory approach, we found that users have incomplete mental models of SPA, leading to different perceptions of where data is being stored, processed, and shared. Users' understanding of the SPA ecosystem is often limited to their household and the SPA vendor at most, even when using third-party skills or managing other smart home devices. This leads to incomplete threat models (few threat agents and types of attacks) and non-technical coping strategies they implement to protect themselves. We also found that users are not making the most of the shopping capabilities of SPA due to security and privacy concerns; and while users perceive SPA as intelligent and capable of learning, they would not like SPA learning everything about them. Based on these findings, we discuss design recommendations.

USENIX Symposium on Usable Privacy and Security (SOUPS) 2019.
August 11–13, 2019, Santa Clara, CA, USA.

1 Introduction

The adoption of smart home personal assistants (SPA) has rapidly increased in the last few years [5]. Estimates suggest that 10% of the world consumers own an SPA [37], and that over 50 million Amazon Echo devices have been sold to date in the US alone [27]. SPA benefit from recent advances in Natural Language Processing to handle a wide range of commands and questions in a playful way, with a name and a gender assigned to the SPA, which encourages users to personify them and therefore interact with them in a human-like manner and be more engaging [32]. SPA are used to shop, stream music, and set timers, alarms and reminders among many others [43].

Despite the numerous benefits and convenience SPA bring, they also raise security and privacy concerns. Prior work, including [12, 17, 22, 28], already highlighted numerous security and privacy issues in general with smart home technologies and in particular with SPA. In addition, very recent research also studied users' privacy concerns with SPA [19, 30], but this research typically centred around privacy and the smart speaker part of the SPA ecosystem. However, smart speakers are just the tip of the iceberg, i.e., an SPA is normally composed of at least a smart speaker such as Amazon Echo *and* a cloud-based voice assistant such as Amazon Alexa. Also, the SPA ecosystem is complex and includes several parties: the SPA provider, multiple third-party providers of skills or actions that SPA can request following users' voice commands (e.g. playing music through Spotify), and multiple providers of other smart home devices (e.g. smart bulbs) being managed through the SPA.

To bridge this gap, in this paper we focus on the following research questions. What are users' perceptions of the SPA architecture and the SPA data ecosystem? What threat models do users have concerning SPA? What mitigation strategies do users use to alleviate risk and other challenges they face?

To answer these research questions, we conducted semi-structured interviews with seventeen current SPA users. Following a grounded theory approach, we interviewed people

who had been using Amazon Echo/Alexa and Google Home-/Assistant, which are the two most used SPA and together dominated circa 87% of the SPA market as of 2017 [39]. We particularly asked about their use of the SPA, how they think SPA process and complete their requests, as well as other data activities like storage, sharing and learning using the four main use cases of SPA: built-in skills (such as setting reminders and alarms), third-party skills (such as Spotify and Uber), managing other smart devices (such as smart bulbs and smart TVs), and shopping. We also elicited users' threat models and the strategies they use to protect themselves when using SPA.

Our contributions include:

- We present users' understanding of SPA's ecosystem, discussing their conceptions and misconceptions about how data is processed, stored, shared and learned by SPA and the actors involved through four main use cases of SPA (built-in skills, third-party skills, managing other smart devices, and shopping). We show that users have a limited understanding of SPA, which leaves them with very inaccurate and at best incomplete mental models of the SPA ecosystem.
- We uncover the lack of trust users have with some of the use cases of SPA, in particular shopping, and how this is hampering adoption of these use cases, providing the reasons we found behind this phenomenon.
- We report the threat models users have of SPA, showing both threat agents and types of attacks users consider possible. We also show the mainly non-technical coping strategies users follow to try to protect themselves.
- and, we present design implications for how SPA might support users' expectations and needs with regards to privacy and security.

2 Background

Smart Home Personal Assistants (SPAs) have a complex architecture [14], as depicted in Figure 1, that usually involves at least a smart speaker (e.g. Amazon Echo, Google Home) and a cloud-based voice personal assistant (e.g. Alexa, Google Assistant). A normal request works as follows, the user utters a request to the smart speaker, which is then processed in the SPA provider's cloud using Natural Language Processing to understand users' speech and intent. Once the intent is identified, the SPA provider delegates the user request to a set of *Skills*[1]. Skills provide users with functions such as the ability to play music, check weather updates, control other smart home devices and shopping. There are currently over 70,000 Alexa skills [1] and 2,000 Google Assistant skills [34]. There are two main types of Skills: Built-in Skills provided by the SPA provider (e.g. Weather updates, Shopping) and Third-party Skills provided by third party developers using the development Skill Kits (e.g. Spotify, Smart Home Devices). Importantly, third-party skills are typically hosted in a remote web service host controlled by the developer of the third-party skill. Finally, any outputs produced by a Skill are sent back to the SPA Provider, which generates a spoken response, which is then push backed to the smart speaker, which plays the response to the user.

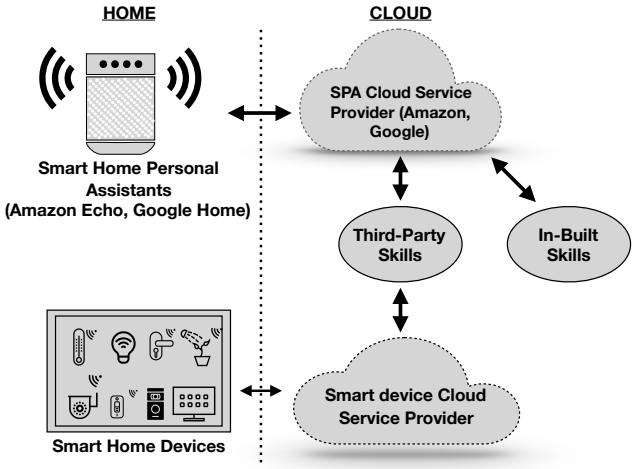

Figure 1: Sketch of the SPA Ecosystem (inspired by [14]).

3 Related Work

In this section, we first discuss research conducted on users' security and privacy perceptions in the Smart Home in general. Then, we discuss research that focused on the security and privacy of SPA in particular.

3.1 Security and Privacy of the Smart Home

Extensive research has been conducted on the security and privacy of smart homes. For example, from a more human factors point of view, prior work studied users' mental models for smart home devices. Zeng et al. [46] conducted semi-structured interviews on fifteen smart home owners examining users' mental models about their device. They found that users who had advanced mental models about their device were those with highly technical level of understanding regarding their smart home system whilst those with intermediate level of mental models showed some level of understanding on how their smart home system works [46]. Similarly Zheng et al. [47] conducted semi-structured interviews with eleven smart home owners to be able to understand their privacy

[1]Note that, for easy of exposition, we adopt Amazon's terminology of Skills, but these may be called differently in other SPA platforms. For instance, in Google Assistant, skills are called *Actions* instead.

perceptions of the devices. Their work highlighted that smart home owners prioritize convenience over privacy and will allow their data to be shared if their perceived benefits outweigh privacy risks. Also users perceived that it is the device manufactures responsibility to protect users privacy. More recently, Emami-Naeini et al. [15] studied security and privacy perceptions of IoT device owners examining their concerns prior and after purchase. Users were asked to rank important factors when they are considering to purchase an IoT device, with security and privacy ranking highly important. They also showed users a security and privacy label prototype aimed at helping them make better security and privacy decisions when purchasing IoT devices [15]. In [36], the authors studied smart home security identifying issues that influence or affect security decisions in the home, e.g., perceived competence, trust and cost were some of the factors identified. Finally, He et al. [24] examined access control specification and authentication in the home IoT, looking at different access controls that can be applied for different tasks depending on the context. While the works above considered the smart home, including SPA, they did not consider the SPA ecosystem in full.

3.2 Security and Privacy of SPA

There has been an increasing amount of research focusing exclusively on SPA security and privacy. One line of research focused on technical attacks and defences. For instance, Haack et al. [22] and Kumar et al. [28] reported vulnerabilities of Amazon Alexa, focusing on the speech recognition ability of SPA (e.g., interpretation errors of user commands exposing the device to outside attacks). In terms of defences, Lei et al. [31] implemented a Virtual Security Button (VSB) which detects the presence of human motion and then prevents unauthorized access. Huan et al. [16] proposed a continuous voice authentication mechanism for SPAs that aims to ensure SPA works solely on commands from a legitimate user. Kepuska et al. [26] proposed a multi-model dialogue system that combines various factors such as; voice, video, head and body movements for secure SPA authentication.

Another line of research focused on human factors of security and privacy in SPA. In particular, previous work studied users' perceptions, including Frutcher and Liccardi [19], who examined users' online reviews of SPA to understand privacy and security concerns. More recently, Lau et al. [30] studied users and non-users reasons for and against adopting SPA. Their findings highlighted that many non-users did not see the benefit in using SPA while users shared privacy risks such as the device listening but would rather trade privacy for convenience. Our work differs from previous work on users' perceptions of SPA security and privacy, as we consider the whole SPA ecosystem, while previous works tended to focus more on the smart speaker part of the SPA only.

4 Methodology

To answer our research questions, we conducted a qualitative study following a semi-structured approach [6] and Grounded Theory [8, 20]. We used a pre-screening process and semi-structured interviews as detailed below. The study was reviewed and approved by King's College London IRB.

4.1 Pilot Study

We created an initial version of the interview script to explore users' perceptions around our main research questions. Before running the full study, we conducted five preliminary interviews. We recruited interviewees internally within our university with the aim of ensuring that the interview questions were easy for interviewees to understand, did not take too long to complete, and would provide insights with regards to our research questions without guiding or biasing the interviewees. With these aims in mind, we conducted and analyzed the preliminary interviews and refined the interview script twice. None of the data collected during the pilot study was used in the final data analysis.

4.2 Recruitment and Screening

We recruited potential participants through Prolific (www.prolific.ac) and internally within King's College London. All potential participants were asked to fill out a screening survey which queried for their demographic information (age, gender, education background, employment status), the SPA and other type of smart devices they own, what they use the SPA for, and how long they have been using the SPA — see Appendix A for the screening questions. The questionnaire took on average 10 minutes to complete and the participants who completed the survey through Prolific were compensated with an average of £1.20.

The screening responses were used to select Amazon Echo or Google Home owners who had been using their SPA for at least one month and had used the device for various tasks such as setting the alarm or reminders, using third-party skills, shopping or managing other smart home devices. Our demographics data also helped us to select participants in a way in which we would maximize demographic coverage. This was done to ensure that selected participants had experience in using the SPA since we wanted to elicit their mental models regarding how SPA work while making sure we had a balanced sample of demographic data. In some cases, the decision was to take everyone who completed the questionnaire in a logical way, but with a particular characteristic, e.g., we invited all valid participants who said they used the device for shopping because of the low number of participants saying they used the device for shopping. Finally, we included some questions designed to rule out participants just pretending to be SPA owners.

The recruitment phase took place between November 2018 and January 2019. The qualified participants were contacted and invited for an interview. Participants were asked to provide their Skype ID. Because this is personally identifiable data, we needed approval from prolific to use such information to recruit participants. We contacted Prolific informing them about our research and the type of data we would be collecting, and our request was approved.

4.3 Participants

From the recruitment and screening process, we received a total of 43 (31 prolific, 9 internal) responses, from which 31 qualified for an interview following the criteria explained above. We contacted all of them, and from the 23 who responded to be available for an interview, we then ordered and prioritized them in order to maximize demographics and SPA usage, until saturation was reached — more details about the methodology and data analysis below. In total, we interviewed 17 participants (13 Prolific and 4 internally). Table 1 summarizes demographic and SPA usage information for all the participants. The interviews were conducted via Skype or in person between January 2019 and February 2019, and participants were rewarded with £10 for completing the interview.

Table 1: Summary of Participants.

	# participants
Gender	
Male	8
Female	9
Age	
18 - 20	2
21 - 25	4
26 - 30	5
31 - 40	3
41 - 50	2
51+	1
Highest Level of Education	
High school/College course	6
Undergraduate	2
Graduate	8
Postgraduate	1
Employment status	
Full time	9
Part-time	3
Unemployed	1
Retired	1
Student	3
Device Type	
Amazon Echo	10
Google Home	7
Period of usage	
1-6 months	5
6-12 months	6
1-2 years	4
2+ years	2
Device use	
Set alarm, reminders and checking the weather	13
Third-party Skills (e.g. Spotify)	11
Managing smarthome devices	6
Shopping	4

4.4 Interview Protocol

Interviews were led and conducted by the lead researcher. Before the interviews, we provided the participants with an information sheet, which explained the purpose of the study. During the interview, the lead researcher introduced themselves and further explained the purpose of the study. We then asked for consent to participate and record audio.

To make participants feel at ease and establish rapport, the interview started with general questions about the participant and their device, we asked them what type of device they owned, what they use it for, how often they used it, how long they have been using it, and whom they were using it with.

The second set of questions focused on asking participants about other smart home devices they own; what devices they owned and if they use their smart assistant to control or communicate with those devices. Then, we asked participants about how they registered and set up their devices. This included questions about voice recognition and purchasing.

To understand and elicit users' mental models about the infrastructure and the data ecosystem, we created four scenarios regarding how the device is used. We would then ask about each scenario depending on the previous answers of the participants, i.e., if they said they had other smart devices they connected the SPA to, then we would ask about the scenario about managing smart home devices. Each scenario was structured as follows. At the beginning of each scenario, we asked participants to think and describe how the SPA worked to complete each request. The second set of questions asked about data storage (including the requests themselves), where data is stored and for how long. The last set of questions focused on whether data was shared and with whom. We describe the scenarios below:

Scenario 1 - Built-in Skills

> In scenario 1, we asked users to think about instances when they asked the device to give them a weather or traffic update. We then asked them to describe how their devices processed their request when using in-built skills. After this, we asked them if such requests are stored, and if yes, where they are stored and for how long. The last set of questions focuses on understanding if data are being used for other purposes than responding to their requests and by whom.

Scenario 2 - Third-party Skills

In Scenario 2, we asked users about third-party skills they used (e.g. Spotify) by asking them to describe how they think the process works regarding how requests are processed and handled. We then asked them whether they think their requests are stored and if so where and for how long. Regarding data sharing, we also asked them if they think their data is shared with third-party skill providers as well as other third-parties such as advertisers.

Scenario 3 - Managing smart home devices

The purpose of scenario 3 was to understand how users perceived SPA's interaction with other smart home devices — e.g., smart bulbs. We asked users to describe how the SPA controls or manages other smart home devices. We began by asking users to think about instances when they controlled other smart home devices with their SPA. Then, we asked them to describe the process to us. We followed asking them if requests are stored. If the user thought these requests were stored, then we continued to ask them where they were stored and for how long. Regarding data sharing, we first asked users if requests were shared with the provider of the smart home device. Then, we asked if SPA's provider (Amazon or Google) together with the smart device provider shared data with other third-party companies.

Scenario 4 - Shopping

In the last scenario, we asked participants to describe how they use the device to shop and how do they think the process works. Similar to other scenarios, we asked them if the device stored their requests including purchase history and for how long. We also asked them if the data was used for other purposes and shared with other third-party companies.

The last set of questions focused on understanding users' threat models concerning the device. Instead of asking participants plainly whether they had security or privacy concerns about using the device, we asked participants what their thoughts were of the SPA capabilities to learn about them based on their interactions with it, who might want to take advantage or exploit the SPA and how, and if they had any concerns about the SPA. Before we concluded the interview, we asked participants about how they protected their devices or mitigated concerns if they mentioned some exploits or other concerns. We provide the final interview script in Appendix B.

4.5 Data Analysis

Following a grounded theory approach [8,20], two researchers independently started the coding process immediately after the first two interviews. Coding was started early to identify interesting codes and categories that could be explored in-depth. The interview scripts were then analyzed through several iterative stages of open, axial and selective coding. When new codes or themes emerged, both researchers met and discussed the new findings and amended the interview script where necessary to explore the new codes or themes in depth. Examples of the codes that emerged very early were: useful, best fit, control, and convenient were prevalent. We discussed and coded these codes under "Useful" as a theme that we defined to denote that users found the device to be useful. New codes stopped emerging after the ninth and tenth transcripts, but we stopped interviewing new participants after number seventeen to confirm we had reached saturation, i.e., to check new codes or themes would not emerge. During the selective coding phase, we ordered and grouped our themes into more broad and abstract groupings to answer our research questions.

5 Findings

This section presents the results of our study. It is structured as follows. It begins by reporting the results in terms of how users use and setup the SPA and the different parts of the ecosystem. Then, we report the different perceptions users have of data processing, storage, sharing and learning across the SPA ecosystem. After this, we focus on the results about one particular use case: shopping, as we found a general lack of trust in SPA shopping capabilities that we study more in-depth considering users who do not shop at all, users who only do part of the shopping process (e.g. shopping lists), and users who do purchase using the SPA. Finally, we report on the threat models users have and the kind of defences and coping strategies they put in place to tackle the threats.

5.1 Device Usage

Participants used SPA for various tasks, all of them falling into the four main use case scenarios:

Built-in Skills. Participants mentioned they used their SPA to complete everyday tasks such as setting an alarm, setting reminders and checking the weather.

Third-party Skills. When asked about third-party skills, participants mentioned using Spotify to listen to music, Uber to call a taxi, Fuel Finder for checking fuel prices, etc.

Managing Other Smart Home devices. Participants also shared using their SPA to manage other smart home devices. In particular, six participants reported controlling other smart home devices. The devices included: smart bulbs, smart TVs and other smart speakers. In addition, some of our participants had tried to connect their SPA to other smart home devices they own but they did not succeed.

Shopping. Four participants use the SPA for shopping. In particular, from those who use SPA for shopping, most of

them use SPA mainly to create shopping lists to later on purchase the items through the website or mobile application, as opposed to purchasing through the SPA. We explore the reasons for this in-depth in Section 5.4.

5.2 SPA Setup

From our 17 participants, 14 reported having set up their SPA while 3 stated that a partner or family member[2] had set up the device for use. All participants who set up the device stated that the setup process was easy and straightforward. They also stated that they used their personal Amazon and Google accounts to set up their devices. These were accounts that users were also using for other personal purposes such as shopping (Amazon accounts), and Android devices (Google/Gmail accounts). Both sets of users reasoned that it was easier and more convenient to use existing accounts than creating new ones, and that they preferred sharing it across the household rather than setting up multiple accounts — *"It is better to share one for convenience sake"* (P2). This is something that, to some extent, one could expect as it was already shown to happen in other home settings [33]. However, other participants reported that they wanted to link the devices with existing accounts and enjoy more of the added functionality and benefits SPA bring to them. As P10 put it *"so that its easy for me to see what's on my calendar"*, i.e., by linking the participant's existing account to their SPA they can set reminders that will sync with their regular calendar system. We found this particularly interesting, as it reinforces the importance of looking at the whole SPA ecosystem not just at the smart speaker placed in households.

5.2.1 Voice Recognition Setup

Although mostly used by the SPA for personalization rather than for security purposes, both Google Home/Assistant and Amazon Echo/Alexa offer voice recognition mechanisms for recognizing the voices of different users, so that they can tell users apart and personalize the interaction with them, named Voice Match [21] and Voice Profiles [2] respectively. In particular, Google users are given the opportunity to configure voice recognition as part of the initial setup process. Six Google users (6/7) setup voice recognition and reported that the device is usually able to distinguish their voices from the others, but with the mechanism being far from perfect, e.g. P12 said *"the times we have tested it seems that it can like but 70% of the time it doesn't seem perfect"*. In contrast, Amazon users are only given the chance to test the speech recognition process (ability to convert spoken words into a text and understand users' intent) as part of the initial setup, but not to configure voice recognition. Voice recognition (in this case Voice Profiles) can be set up at any point but always after the initial setup and as a separate process. Only 2 out of the 10 Amazon users reported completing voice recognition. Most users who did not set it up did not even know that this mechanism actually exists. Interestingly enough, some of those who did not complete voice recognition seemed not to understand or differentiate between speech recognition and voice recognition, and they would confuse them, thinking the SPA can distinguish between people without having set voice recognition. For those who understood the difference, they explained that speech recognition was a feature that allowed the device to recognize speech and change it to text while voice recognition involved the device being able to tell who is talking. When asked how the process works, they revealed that the device has an AI system which compared voices to distinguish between users. While these group of users reported that voice recognition is used to distinguish between users, some said it was for recognizing different accents (actually meaning speech recognition).

[2]Note here that we did not get into the tensions between those setting the devices and other household members, as this was already studied in-depth, including SPA too, in [46].

5.2.2 Third-party Skills Setup

Some third-party skills need to be setup either in terms of the permissions they need to access, e.g. smart speaker country and postcode for the case of Fuel Finder, or to link them to other online accounts to provide the functionality required, e.g., playing music through Spotify. We asked participants to describe the process of setting up the third-party skills they use. In some cases, this already started to shed light about their mental models. To setup the skill users share configuring their SPA to the skill they want to use. For example, P11 said: *"directly connected to my spotify account so it directly logs in to spotify and play music"*.

5.2.3 Connecting to Smart Home Devices

Managing other smart home devices through an SPA obviously requires connecting the SPA to the device. We asked participants to describe the process of connecting their SPA to their other smart home devices. They mentioned downloading the other smart home device mobile application and configuring it to their SPA, for example P3 stated *"I have the app on my phone so I use that to manage my activities between Alexa and the lamp"*. Other participants shared negative setup experiences, with some of them unable to connect their smart home devices to the SPA, with P9 stating *"I was unable to configure my smart TV, google home can't find the device"*.

5.2.4 Shopping Setup

Both Amazon Echo/Alexa and Google Home/Assistant support shopping lists by default, so users can just create lists and add items to buy. In terms of actually completing a purchase, Amazon allows users to optionally create a 4-digit pin code to be used when purchasing online. In particular, one of our

participants had set the code. Others who did not use their SPA for any shopping activities simply did not have the voice purchasing code setup and had it disabled. It should be noted that if a user has Amazon Voice Profiles (voice recognition) setup, they need to setup the pin code for purchasing but do not need to say it every time they want to complete a purchase.

5.3 Perceptions of SPA's Ecosystem

In order to explore what perceptions users had of the ecosystem, we considered all the four main use case scenarios and asked about different information-related activities (data processing, data storage, data sharing, data learning) and how they thought these activities where being conducted and where.

5.3.1 Data Processing

In general, our analysis shows that most SPA users believe that data collected by the device is processed locally in the device, though a few reported that the device needed to be connected to the Internet to work. Others explained that their requests are processed remotely and relayed back to the device. We explain this in detail below per type of use case, as there were some interesting differences worth mentioning across them.

Built-in Skills. When asked to describe how the device processes and fulfills requests like weather updates, 10 out of the 13 participants that used built-in skills explained that the device locally processes these commands and respond to the user. For instance, one user described the device as a small brain, implying that the device listens and process commands before responding to the user. We also found a few participants who believe that the device communicates with the SPA provider to process commands and then responds to the user, but in many of these cases, this was because they thought the SPA connected with an online source of information to process requests. For instance, one participant mentioned that the device connected to the Google website for weather updates. P9 shared this *"... with the weather. I believe it comes from the Google site from their weather service"*.

Third-party Skills. We observed that 10 out of the 11 participants who use third-party skills do not consider the third-party skills providers when describing how SPA process their request when a third-party skill is involved. While some users reported that data is sent to the SPA provider for processing, they did not mention any communication between the SPA provider and the third-party skill provider. This contrasts sharply with the very few participants who had a better understanding, though still incomplete and inaccurate, of how the process works when the SPA uses third-party skills. For instance, P8 stated *"well Alexa when I say I want to play a song she'll then connect to Spotify and search through the catalogue I guess then play the song"*.

Managing Smart devices. We found that 5/6 participants believe that the smart speaker and other smart home devices communicate directly without involving other elements of the SPA architecture. For instance, when switching the smart lights on, they believe that the device communicates directly with the lights, implying that both the SPA provider and the smart light provider are not involved in any way. Some participants think these devices communicate through the mobile app (i.e., other smart home device's mobile app) installed in their mobile phones. For example, P2 said: *"basically Google Home talks to the light bulb via the mobile app installed, and they are connected via the network so I will say OK Google turn the light off/On and it will send the request to the application that controls the Philips light"*.

Shopping. We found users shopping using the SPA talked about voice purchasing much in the same way as they would do for normal online purchases. For instance P5 said *"I just ask Alexa to add items to my basket"*. They also mentioned the SPA provider as being somehow involved in the process as the market/account they were buying from, e.g. P13 said *"once i ask alexa to add item to my basket she updates it on my Amazon account"*. While most of our participants did not complete the purchasing process using the device, we asked all users for their views concerning voice purchasing. The majority (13 out of the 17), reported having not thought about the process, but we observed that, similarly to those who use SPA to shop, their current online shopping practices influence their understanding of the voice purchasing process. They think about how they would select an item to buy, choose the method of payment and confirm the order.

5.3.2 Data Storage

In general, most users believe that their voice recordings, the history log and shopping history are all stored by the SPA provider. These users think this information is kept for building a profile about them, i.e. to understand their behaviour and interests. Regarding where data is stored, half of our interviewees believed that data collected by the SPA is stored locally in the smart speaker, while others reported that data is stored either in the cloud owned by the SPA provider, or both in the smart speaker and the cloud. One user stated that data is not stored at all since there is so much data to store. All our interviewees informed us that they do not know how data is stored and how long the provider keeps it. We further explain users' perceptions of how data is stored below depending on the use case.

Built-in Skills. When using the device to complete everyday tasks like setting reminders, asking questions and requesting updates (e.g., weather), most participants believed that data is stored to learn more about them and personalize the SPA experience. Half of the participants believed that these data were stored locally in the device. One user stated such data was not stored at all since there were many data to store and the provider would not be able to handle it all. Another user mentioned that data (i.e., history) were stored in the mo-

bile app.

Third-party Skills. Users who were using their SPAs with other third-party applications reported that their requests and history logs (e.g., playlists) were only stored by the SPA provider not mentioning the third-party provider. We observed that most users of third-party skills (9/11) do not mention their third-party skills providers storing any data.

Managing Smart Devices. Most users who use their SPA to manage other smart home devices (5/6) reported that their commands directed to the smart home devices were stored, but they only assumed that it was to personalize and improve their SPAs. In terms of where or who stores the information, none of the participants mentioned any of the providers of the smart home devices they were managing through the SPA. They seem to only believe that their smart speaker stores all data locally in this case.

Shopping. When using voice purchasing, users believe that their shopping lists and history are stored by the SPA provider. They reasoned that this is done to understand their shopping interests and behaviour. While the majority believed that this is for advertising purposes, some believed this is for improving the SPA. Regarding deletion of data, some participants stated that shopping history is immediately deleted from the device after shopping.

5.3.3 Data sharing

While the usage of the SPA includes data being shared by the SPA provider and other different vendors or third-parties, we observed that users' perception of how data is handled and shared is mostly based on the stories of data misuse they know from other domains. For instance, users believe data is shared with data brokers and third-parties who are interested in influencing their behaviour, as P3 explained: *" ...so they would to try and influence users purchasing decisions"*. Other participants alluded to the Facebook and Cambridge Data Analytica case [45] and stated that they did not know with whom their data is being shared but believe it was being shared with other companies *P4 "they could give it to third-party people to target certain adverts to the user"*. However, some users reported that precisely because of recent data misuse incidents, they trusted their SPA providers not to share data with other parties.

In terms of the wider ecosystem, none of the participants who used the SPA with third-party skills (e.g., Uber) or with other smart devices (e.g., Phillips bulbs) mentioned data being accessible to these third parties (e.g. Uber or Phillips), let alone with whom these third parties might be sharing the data they gather. That is, no users mention the fact that, because they may have access to users' data because of how the SPA ecosystem work, that they could too share that data with others, not just the SPA provider. In terms of the specific data that participants believe is being shared, our participants informed us that their usage statistics, shopping habits and play-lists are being shared with other parties like advertisers.

5.3.4 Data Learning

We also asked participants whether their SPA were capable of learning things about them based on their interaction or usage. In general, participants perceive SPA as intelligent and having the ability to learn new things about them without they telling them to the SPA. They describe them as a brain or having a memory to process and remember certain things about them. Others describe their device as an Artificial Intelligence (AI) system. In a similar way to processing and storing data, users seemed to attribute all the learning capabilities to the smart speaker as opposed to other parts of the SPA ecosystem involved in it, which they did not mention. They tended to personify the smart speaker and say it was intelligent.

Regarding how the device learns about them, 13/17 participants said the SPA analyses their usage patterns (i.e., questions, play-lists, history logs and shopping lists) to learn about their likes and dislikes. Our analysis shows that users believe their SPA are capable of learning about their shopping habits, their favourite music and radio stations, routines and its users. They also believe that the device uses what it learns about them to tailor adverts for them, serve them well, to influence their decisions and recommend better things to them (e.g., more music from their favourite artist), P17 *"It picks up adverts for example on my android phone I get adverts related to what I have asked my Google Home so it shows that element of the device listening"*. While P7 explained: *"I would probably imagine it stores your information and it [then] begins to predict through [the data] I would assume... some sort of like a pattern, therefore, it would [then] start to tailor things to people that fit that [particular] pattern."*

Some users have mixed attitudes toward the device being able to learn things about them; some perceive this as a negative trait while others see it as a useful feature, with some perceiving both depending on the context — reasoning very similarly to what well-known theories like Contextual Integrity [35] aim to explain. For instance, some users stated that the device being able to learn and know certain things about them (e.g., morning routine – favourite music, weather, traffic and news updates) is a good thing as it could simplify their life. However, they explained that it is not pleasant for the device to know sensitive things about them, for instance, health symptoms.

In general, users (including those that perceive learning as a good thing) find the idea that the device can learn about them being creepy, scary and invasive, sometimes because they could never tell when the device is doing the actual learning. P9 explained: *"In a way, it is good for it to give you suggestions. But, at the same time, it is scary because if you think about it, if it's learning things you are doing it is quite sinister. At the moment I am happy with it, but it does make me think about what information it can learn about me... what*

profile it can build without me realising".

5.4 Shopping

In our initial interviews, it quickly became apparent that one of the use cases we where considering, shopping, was actually worth studying more in-depth because of a seemingly low adoption by current SPA users. Both Amazon Echo and Google Home devices give users the opportunity to shop online. Therefore, we asked all participants about their shopping experience, the challenges and the concerns they have while using the device to shop. We aimed to understand how users view the process of shopping, from the moment they make the shopping list to the point of payment. We particularly sought to understand any differences in perception/use among those who do not use SPA to shop at all, those who use the SPA to aid their shopping even if not purchasing through the SPA (e.g. just using shopping lists), and those who actually use the SPA to purchase items.

Users view voice purchasing as a convenient way of shopping, with some tasks such as creating shopping lists and paying for goods faster than with other systems. We found that most of our participants (8/10 Echo users) had not set up voice purchasing code because they were not using voice purchasing features on their SPA. In general, most participants (7/10) told us that the voice purchasing code is a useful feature of the device and adds an extra layer of security. However, further analysis showed that most users are concerned that other people around the house (or neighbours) could hear the code and use it maliciously.

Below we summarize users' main struggles and concerns about using the SPA to shop. Mainly, we observed trust, or more specifically the *lack of trust*, emerging very strongly as a theme across different dimensions: products (visibility, comparisons, and mistakes), vendors, security of the connection, and privacy of the orders. In particular:

Product visibility. When we asked our interviewees their thoughts on using the device for shopping, 10 out of 17 participants stated their biggest concern not being able to see the product they want to purchase.

For instance, P12 said: *"I [am] probably kind of against it, cause I will need [a] screen to see what I am buying, I need a lot of confirmation; how the products are and what I am buying, so a visual thing. So just using voice assistant I don't think I would ever do that.".*

Product comparisons. Some users expressed the difficulties of comparing products when shopping using the device. Some Amazon Echo users stressed that Alexa did not give them a full description of the product but just the name and the amount. Other Amazon Echo users noted that one could not get the reviews of the product. Users also raised some concerns about fake products, that using the device one may end up ordering a fake product. For instance, P1 explained why he is not using voice purchasing: *"...erm only because I am aware of scams and fake products on Amazon, I would like to see what I am buying first."*

Product mistakes. We found that some users were concerned about buying the wrong item because, sometimes, while they are creating the shopping list, the device gets the wrong item. In fact, two participants reported an unsuccessful attempt of shopping using the device as they mentioned the wrong items were added into their shopping basket. For instance, P5 said *"I just ask Alexa to add items to my basket and it does, but often it adds the wrong items..."*.

Number and trustworthiness of vendors. Users expressed that they have a limited number of vendors to buy from than when shopping online. For instance, one Amazon user argued that they are limited when using the device because they cannot buy from other outlets. However, some Google Home users thought being connected to a single vendor (like Amazon Echo users) guarantees security as the user is just connected to a well-known and trusted outlet. Nonetheless, some Google Home users informed that the number of vendors is limited and there is a chance of not finding what they want. Other Google Home users stated that it is difficult to choose which vendor to use.

Secure connection. Some users, mainly those who had not set up voice purchasing expressed their concerns over secure payment and connection during shopping. They stated that it was challenging to confirm whether they are connected to the right vendor or the payment process is secure. P9 said, *"... I don't know if the payment is secured or it's going through an encrypted site as a basic example, I like to see something on a screen rather than doing it on an automated home system."* Moreover, some further informed us that there are no visual cues to help them feel they are secure.

People hearing orders and/or code. Some users who were not using the device to shop highlighted some privacy concerns of other people being able to hear what they are ordering, for instance, P14 stated *"people around you can easily hear your purchasing code which isn't safe if you think about it"*. Others said its easy for other members of the family or neighbours to hear what they are ordering and that can be unpleasant at times. They also mentioned concerns about others hearing the voice purchasing code and using it without their permission.

The above struggles and concerns make users utilize a number of strategies in order to minimize the concerns. Those include:

Completing the order through the app. To avoid buying wrong items, some users stated that they use the device to create shopping lists but always confirm their orders before paying through mobile apps or website. For example, P5 said: *"I just ask Alexa to add items to my basket and it does ... and [then] I have to go to the app to make the purchase."* Most users who used this strategy mentioned that the device is good for making shopping lists but not ideal for shopping especially when product details matter.

Disabling voice purchasing. Most participants mentioned that they decided not to enable voice purchasing because they do not trust it. These were users who earlier revealed that they were not sure how secure the device is when shopping.

Shopping through other platforms. Some users explained that they still prefer to shop using their apps or the web. They explained that shopping using other platforms gave them the opportunity to find better deals and get the right items. Some users further explained that these other platforms are trustworthy and have been using them for some time. For example, P10 noted: *"I would say the device is great for other things, but in terms of shopping, it is useful to add things to your basket, but I would say its better to buy through the website or app, so you know its safe and secure."*.

5.5 Threat Models and Coping Strategies

To understand users' privacy and security concerns regarding owning and using SPAs, we asked users if they thought their devices could be exploited maliciously or if their data could be at risk while using the device. We were also interested in the threat agents – actors who might be interested in such attempts. Considering the size of the ecosystem and the number of stakeholders involved, these questions aimed at getting participants to describe the threats and the attacks that SPAs might be subjected to. After these questions, we wanted to know what users do to protect themselves from these threats and attacks.

All of our participants reported that their SPA could be exploited. They described how different threat agents could attack the device. In general, we observed many gaps in their threat models; users consider few threat agents and exclude the people they share the device with. Also, they do not consider malicious skills or SPA providers. Users are mostly worried about unwanted listening from the device. They reported not knowing how to protect themselves or their devices from technical attacks but shared various non-technical solutions they develop to protect themselves.

5.5.1 Threat Agents

While some of the users explained that anyone could *hack* the SPA, the most common threat agents that users discussed were: hackers, government agencies and data brokers (advertisers). Many of our participants used words like "criminals" and "fraudsters" to describe potential threat agents. We grouped all these under the theme "Hackers". Users gave various reasons, i.e. *motivations*, to why these threat agents would be interested in attacking the SPA. They mentioned that hackers (and fraudsters) would be interested in targeting SPA for financial gain, to get personal data which they can then sell and for blackmailing purposes; government might do it for spying on users and influencing their decisions; and advertisers would do it for understanding users' usage behavior and use that for marketing purposes. We also found that participants who mentioned advertisers highlighted that data generated by users is considered important, and everyone is interested in it. However, most users who mentioned "fraudsters" and "criminals" linked them to financial gains. For instance, P4 stated *"... with the shopping feature [available], potential people who want to steal money of you can target it... because your credit card is stored so I would say fraudsters"*.

Despite recent news, e.g., Amazon releasing a user's Amazon echo recordings to another user [44], users do not normally consider the SPA providers or providers who have access to the data, e.g. third-party skill providers, as threat agents. Also, no one mentioned other household members as a potential source of problems. However, studies suggest that smart home devices are weaponized within the family [23]. The only hint towards this was a few participants who mentioned the problem of other household members and neighboors overhearing the voice purchasing code.

5.5.2 Attack Types

We found that while users' threat models consist of different threat agents, many users struggled to describe attacks that SPAs can suffer. The most prevalent attack mentioned by our users is unwanted listening. All our users raised this as a concern and mentioned different threat agents hacking the device to listen and spy on them. Some users shared advanced attacks such as attacking the network the device is connected to and hijacking the commands, but still attacks did not normally correspond to the real attack surface of SPA (see Section 3). For example P17 shared *"They are connected to the network so they have IP and storage so they can become part of a botnet"*. Also, they hardly related to any parts of the SPA ecosystem but the smart speaker — e.g. participants did not consider malicious skills [29].

5.5.3 Coping Strategies

We found that users do not take any technical solutions to deter threats or protect themselves. We now discuss the strategies they follow to protect themselves below.

Unable to protect themselves. Many participants reported that it is difficult to protect the device because they do not know what attacks might affect their devices. P4 said: *"With these sorts of things, I don't really know if there is a way of protecting yourself..."*. P1 further explained: *"With my PC and phone, I have an anti-virus [installed], but I don't know how you could protect a speaker..."*. This is remarkable, as it shows many users, even if they might do something to better protect themselves, simply cannot do it because they do not know what to do.

Not enabling certain features. Users reported that they disable (or do not set up) some features and functionality of

the SPA to minimize or avoid being at risk. One example of this, as mentioned above, is disabling voice purchasing to avoid risks associated with shopping. This means many users are just restricting themselves in terms of the SPA capabilities they could be using. P10 said: *"Somehow yes, I would limit the things I use it for like I wouldn't use it for purchasing at all I'll stick to shopping lists."*

Using other devices. Some users reported that they use other devices to complete specific tasks in order to minimize what the device knows about them. For instance, P9 said: *"...checking the weather would be ok, but I would be concerned, for example, if I wanted to find out about a certain illness a family member has, I wouldn't do it through Google home...I would use the computer [be]cause I don't want that to be stored [in the SPA]"*. Another example of this is that, as mentioned above, participants tend to complete purchases using other devices like mobile apps after having created a shopping list with the SPA. Again, this means that users are simply not using the SPA for tasks it is capable of doing.

Turning off or muting the smart speaker. We found that some users switch off their SPAs to stop them from listening. They turn the device off when they are sleeping, having private conversations and when they are not home to avoid unauthorized people using them. P9 explained: *"...I would turn it off when we are not in the house so people can't access it when we are not in."* This finding confirms what was also found in [30], where they asked about whether users used the muting button of smart speakers, which in turn revealed that many users were turning off the speakers altogether.

6 Discussion

We now discuss our findings, their implications, and some recommendations.

6.1 More than Smart Speakers

The majority of users see the smart speakers as the place from the whole SPA ecosystem where most of the data processing, storage and learning happens. For instance, when asked to describe how the SPA process and fulfill requests like weather updates, the majority of our participants explained that the device locally processes these commands and respond to the user, mainly limiting the SPA to the smart speaker, which would in turn be some kind of a small brain. This shows that most users have a very simple and inaccurate mental model of the SPA ecosystem. Even those who actually recognize that the SPA needs to search for and find information online do have incomplete mental models. Very few participants clearly involved the SPA provider in the processing, storage, sharing and learning capabilities of SPA, let alone other important actors in the ecosystem like the third-party skills providers and the vendors of smart home devices they manage through the SPA. Therefore, better awareness and transparency mechanisms may help users understand how SPA operate, not necessarily from a technical point of view, but just enough to understand the implications in terms of their data. Awareness and transparency mechanisms, however, need to be engineered carefully, to avoid these mechanisms becoming a lot of information to digest, which may intimidate and/or become a burden on the users, ultimately ending up of not much use. In fact, some participants actually complained about SPA privacy policies and terms of service not being clear enough for them to understand how their data is handled, something that one would expect as it is the case in other domains [13, 25, 38, 40]. Recent research suggested that privacy notices should be relevant, actionable and understandable [41]. In particular, the authors identify four main dimensions to consider when designing to provide notice: timing, when should a notice be presented; channel, how should the notice be delivered; modality, how the information should be conveyed; and control, how choice options are integrated into the notice. Research on SPA notices exploring those dimensions would be really interesting, particularly as conveying notice across the SPA ecosystem, considering its complexity and the actors involved, is non-trivial.

6.2 What do I do to protect myself?

Having better transparency mechanisms that help improve users' mental models of the SPA ecosystem may not necessarily mean that users are able to protect themselves better. Although most participants were clearly unaware of the potential threats, which could mean that they underestimate the security and privacy risks of SPA, and one might be tempted to attribute this to the inaccurate and incomplete mental models users have, one of the main problems we encountered is that most users simply did not know what they could do to protect themselves when using SPA. This actually leads to a situation whereby users minimize the use they make of the SPA to just the cases they think (whether actually right or wrong) are less dangerous. If users are to make the most of SPA, we definitely need more usable security and privacy mechanisms that seemingly integrate with the SPA ecosystem, together with the awareness and transparency mechanisms already mentioned, which in turn may help increase users' trust on SPA.

AI to personalize security/privacy. The first example of potential mechanisms to explore would be those that could leverage the AI capabilities SPA have to personalize the experience to users, so that they would be used to personalize users' security and privacy experience. This would contribute to the cases we found participants felt SPA learning is a good thing. In this way, recent research already suggested variables to consider for permissions within a household [24]. This, together with permissions across the entire SPA ecosystem considering the actors involved, could be the basis for SPA to learn what are the kind of contextual social norms that apply

for particular users and households to govern data processing, storage, sharing, and learning based on the context to help users manage and control their data across the SPA ecosystem. In fact, the feasibility of learning contextual social norms was already shown in other domains [7, 11, 18, 42], and more recently in generic smart homes [4], but this still needs to be considered in the context of the whole SPA ecosystem.

Voice recognition for usability and as building-block. We observed that when the voice recognition setup process is included in the initial SPA setup as with Google Home/Assistant, many more participants seemed to configure it and, actually, they found the process easy and straightforward. In contrast, voice recognition in Amazon Echo/Alexa is not part of the initial setup, and the vast majority of users had not tried to set voice recognition, with some of them not even knowing the mechanism exists. While voice recognition may still not be a mature-enough authentication mechanism in SPA, as it has been shown to be vulnerable to attacks such as spoofing through replay attacks [9], there is indeed ongoing research to make it more secure [16]. The good news is that, from a usability point of view, this looks like an interesting research line, because of the aforementioned proportion of participants who went for voice recognition when they knew about it and were given the chance to set this up at the initial setup stage. Voice recognition could also be the basis for other security mechanisms or to increase trust in some SPA use cases such as shopping, as explained next.

6.3 Trusted Shopping

We found a lack of trust from users when shopping using the SPA. While some participants found it useful and convenient to use some of the SPA's shopping capabilities such as shopping lists, participants would not normally purchase the items through SPA. We identified that the main cause of this was that users did not trust the products, the vendors, and the process, including the security of the connections and whether other people might be able to overhear their purchases and purchase codes. These trust issues need to be addressed in order to foster purchases through SPA, even more if we look towards a future where we will delegate more and more tasks to SPA [10]. Research on the particular mechanisms to make purchasing through SPA more trustworthy seems like an exciting line of future research. For instance, in terms of products and vendors, novel ways for an SPA to somehow provide more verbal information about the products and the vendors, such as product reviews or vendors' reputation, would need to be engineered in a usable way. Also, this type of assurances might need to be complemented with other modalities, something that may be easier with the new generation of multi-modal smart speakers, such as the new 2nd generation of Amazon echo, which includes a screen [3] users could use to check the products in the shopping list to purchase all in one place, with the SPA quickly ordering the items as soon as the user confirms verbally. Also, and as mentioned above, having voice recognition from the beginning would make it so the voice purchasing codes needed in Amazon would not need to be repeated in each purchase (as it is actually the case [2]), also mitigating the concerns some users had in terms of others overhearing the code and using it.

6.4 Limitations

The methodology used was mostly qualitative and exploratory in nature, therefore the hypotheses we formulated based on our findings, emerging themes and discussion coming from the grounded-theoretic analysis, would obviously need to be tested in a follow-up confirmatory study to assess their validity and generalizability. We focused on current SPA users, so we could explore the ecosystem and the parts they understand or use more/less and why, e.g. lack of trust regarding SPA shopping, and because previous work [30] had already looked at users and non-users of SPA to study reasons for adoption. Finally, the interviews were conducted Jan-Feb 2019 after some major news, including Amazon sending thousands of recordings to the wrong user [44], but before the most recent news regarding Alexa recordings being analysed by humans. While this might alter mental models regarding the SPA provider, sometimes mentioned by participants, it might not regarding third-party skills or other third-parties, who were not in the news and hardly mentioned by participants. Nevertheless, understanding how such news could alter users' mental models, particularly in terms of the SPA provider, should be studied.

7 Conclusion

This paper reports our study of users' perceptions of the SPA ecosystem through semi-structured interviews around four main use cases of SPA (built-in skills, third-party skills, managing other smart home devices, and shopping). We uncovered users' misunderstanding of SPA ecosystem, with most users showing a very limited conception of SPA and inaccurate and incomplete mental models of the SPA ecosystem and related data activities (processing, storing, sharing, and learning). We also uncovered the lack of trust users have with some of the use cases of SPA, and how this is hampering adoption particularly of purchasing through the SPA, with users not having enough information to assess their trust in the products, the vendors, and the process of voice purchases. In addition, we reported the threat models users have of SPA, showing both threat agents and types of attacks users consider possible. We also show the mainly non-technical coping strategies users follow to try to protect themselves. Finally, we presented design implications for how SPA might support users' expectations and needs with regards to privacy and security, including researching on mechanisms that help increase awareness, transparency, control, and trust across the SPA ecosystem.

References

[1] The alexa skill store for france is a fast growing land of opportunity, https://voicebot.ai/2018/11/03/the-alexa-skill-store-for-france-is-a-fast-growing-land-of-opportunity/, 2018.

[2] Amazon. *About Alexa Voice Profiles*, 2019 (accessed February 22, 2019). https://www.amazon.com/gp/help/customer/display.html?nodeId=202199440.

[3] Amazon. *All-new Echo Show (2nd Gen) – Premium sound and a vibrant 10.1" HD screen*, 2019 (accessed February 22, 2019). https://www.amazon.com/All-new-Echo-Show-2nd-Gen/dp/B077SXWSRP.

[4] N. Apthorpe, Y. Shvartzshnaider, A. Mathur, D. Reisman, and N. Feamster. Discovering smart home internet of things privacy norms using contextual integrity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(2):59, 2018.

[5] S. Bay. Ai assistants are poised for major growth in 2018. 2018.

[6] Bryman. *Social research methods*. Oxford university press, 2015.

[7] G. Calikli, M. Law, A. K. Bandara, A. Russo, L. Dickens, B. A. Price, A. Stuart, M. Levine, and B. Nuseibeh. Privacy dynamics: Learning privacy norms for social software. In *Proceedings of the 11th International Symposium on Software Engineering for Adaptive and Self-Managing Systems*, pages 47–56. ACM, 2016.

[8] K. Charmaz. *Constructing grounded theory*. Sage, 2014.

[9] S. Chen, K. Ren, S. Piao, C. Wang, Q. Wang, J. Weng, L. Su, and A. Mohaisen. You can hear but you cannot steal:defending against voice impersonation attacks on smartphones. *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, 2017.

[10] P. Cohen, A. Cheyer, E. Horvitz, R. El Kaliouby, and S. Whittaker. On the future of personal assistants. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 1032–1037. ACM, 2016.

[11] N. Criado and J. M. Such. Implicit contextual integrity in online social networks. *Information Sciences*, 325:48–69, 2015.

[12] T. Denning, T. Kohno, and H. M. Levy. Computer security and the modern home. *Communications of the ACM*, 56(1):94–103, 2013.

[13] J. B. Earp, A. I. Anton, L. Aiman-Smith, and W. H. Stufflebeam. Examining internet privacy policies within the context of user privacy values. *IEEE Transactions on Engineering Management*, 52(2):227–237, May 2005.

[14] J. S. Edu, J. M. Such, and G. Suarez-Tangil. Smart Home Personal Assistants: A Security and Privacy Review. *arXiv eprint arXiv:1903.05593*, 2019.

[15] P. Emami-Naeini, H. Dixon, Y. Agarwal, and L. F. Cranor. Exploring how privacy and security factor into iot device purchase behavior. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 534. ACM, 2019.

[16] H. Feng, K. Fawaz, and K. G. Shin. Continuous authentication for voice assistants. In *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*, pages 343–355. ACM, 2017.

[17] E. Fernandes, J. Jung, and A. Prakash. Security analysis of emerging smart home applications. In *Security and Privacy (SP), 2016 IEEE Symposium on*, pages 636–654. IEEE, 2016.

[18] R. Fogues, P. K. Murukannaiah, J. M. Such, and M. P. Singh. Sharing policies in multiuser privacy scenarios: Incorporating context, preferences, and arguments in decision making. *ACM TOCHI*, 24(1):5, 2017.

[19] N. Fruchter and I. Liccardi. Consumer attitudes towards privacy and security in home assistants. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, page LBW050. ACM, 2018.

[20] B. G. Glaser and A. L. Strauss. *The discovery of grounded theory: Strategies for qualitative research*. Transaction publishers, 2009.

[21] Google. Set up multiple users for your speaker or smart display. https://support.google.com/assistant/answer/9071681, 2017. Last accessed 20-February-2018.

[22] W. Haack, M. Severance, M. Wallace, and J. Wohlwend. Security analysis of the amazon echo. *Allen Institute for Artificial Intelligence*, 2017.

[23] M. Hansen and B. Hauge. Scripting, control, and privacy in domestic smart grid technologies: Insights from a danish pilot study. *Energy research & social science*, 25:112–123, 2017.

[24] W. He, M. Golla, R. Padhi, J. Ofek, M. Dürmuth, E. Fernandes, and B. Ur. Rethinking access control and authentication for the home internet of things (iot). In *27th USENIX Security Symposium (USENIX Security 18). USENIX Association, Baltimore, MD*, 2018.

[25] P. G. Kelley, J. Bresee, L. F. Cranor, and R. W. Reeder. A nutrition label for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, page 4. ACM, 2009.

[26] V. Kepuska and G. Bohouta. Next-generation of virtual personal assistants (microsoft cortana, apple siri, amazon alexa and google home). In *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 99–103. IEEE, 2018.

[27] B. Kinsella. The Information Says Alexa Struggles with Voice Commerce But Has 50 Million Devices Sold. https://voicebot.ai/2018/08/06/the-information-says-alexa-struggles-with-\voice-commerce-but-pass, 2018. Last accessed 28-February-2019.

[28] D. Kumar, R. Paccagnella, P. Murley, E. Hennenfent, J. Mason, A. Bates, and M. Bailey. Skill squatting attacks on amazon alexa. In *27th {USENIX} Security Symposium ({USENIX} Security 18)*, pages 33–47. USENIX, 2018.

[29] D. Kumar, R. Paccagnella, P. Murley, E. Hennenfent, J. Mason, A. Bates, and M. Bailey. Skill squatting attacks on amazon alexa. In *27th USENIX Security Symposium (USENIX Security 18)*, pages 33–47, Baltimore, MD, 2018. USENIX Association.

[30] J. Lau, B. Zimmerman, and F. Schaub. Alexa, are you listening?: Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):102, 2018.

[31] X. Lei, G.-H. Tu, A. X. Liu, C.-Y. Li, and T. Xie. The insecurity of home digital voice assistants-amazon alexa as a case study. *arXiv preprint arXiv:1712.03327*, 2017.

[32] E. Luger and A. Sellen. "like having a really bad pa": The gulf between user expectation and experience of conversational agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 5286–5297, New York, NY, USA, 2016. ACM.

[33] T. Matthews, K. Liao, A. Turner, M. Berkovich, R. Reeder, and S. Consolvo. She'll just grab any device that's closer: A study of everyday device & account sharing in households. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 5921–5932. ACM, 2016.

[34] A. Mutcher. Google assistant app total reaches nearly 2400 thats not real number really 1719, 2019. Last accessed 22-Feb-19.

[35] H. Nissenbaum. Privacy as contextual integrity. *Wash. L. Rev.*, 79:119, 2004.

[36] N. Nthala and I. Flechais. Informal support networks: an investigation into home data security practices. In *Fourteenth Symposium on Usable Privacy and Security ({SOUPS} 2018)*, pages 63–82, 2018.

[37] OVUM. Virtual digital assistants to overtake world population by 2021. 2017.

[38] I. Pollach. What's wrong with online privacy policies? *Communications of the ACM*, 50(9):103–108, 2007.

[39] S. T. S. Portal. *Worldwide intelligent/digital assistant market share in 2017 and 2020, by product*, 2019 (accessed Feb 22, 2019). https://www.statista.com/statistics/789633/worldwide-digital-assistant-market-share/.

[40] K. M. Ramokapane, A. C. Mazeli, and A. Rashid. Skip, skip, skip, accept!!!: A study on the usability of smartphone manufacturer provided default features and user privacy. *Proceedings on Privacy Enhancing Technologies*, 2019(2):209–227, 2019.

[41] F. Schaub, R. Balebako, and L. F. Cranor. Designing effective privacy notices and controls. *IEEE Internet Computing*, 21(3):70–77, May 2017.

[42] Y. Shvartzshnaider, S. Tong, T. Wies, P. Kift, H. Nissenbaum, L. Subramanian, and P. Mittal. Learning privacy expectations by crowdsourcing contextual informational norms. In *Fourth AAAI Conference on Human Computation and Crowdsourcing*, 2016.

[43] M. Singleton. Alexa can now set reminders for you. https://www.theverge.com/circuitbreaker/2017/6/1/15724474/alexa-echo-amazon-reminders-named-timers, 2017. Last accessed 28-February-2019.

[44] N. Statt. *Amazon sent 1,700 Alexa voice recordings to the wrong user following data request*, 2018 (accessed Jan 22, 2019). https://www.theverge.com/2018/12/20/18150531/amazon-alexa-voice-recordings-wrong-user-\gdpr-privacy-ai.

[45] A. Valdez. *Everything You Need to Know About Facebook and Cambridge Analytica*, 2018 (accessed Jan 22, 2019). https://www.wired.com/story/wired-facebook-cambridge-analytica-coverage/.

[46] E. Zeng, S. Mare, and F. Roesner. End user security and privacy concerns with smart homes. In *Thirteenth Symposium on Usable Privacy and Security ({SOUPS} 2017)*, pages 65–80, 2017.

[47] S. Zheng, M. Chetty, and N. Feamster. User perceptions of privacy in smart homes. *arXiv preprint arXiv:1802.08182*, 2018.

A Screening Questions

1. Which device do you own?
   Amazon Echo
   Google Home
   Apple Homepod
   Microsoft Cortana
   2. How long have you been using the device?
   3. How many people within your household use the device?
   4. Which of the following voice commands are used to awaken Amazon's personal assistant?
   "Alexa"
   "Computer"
   "Hey Amazon"
   "I don't own this device"
   5. Which of the following voice commands are used to awaken Google personal assistants?
   "Hey Google"
   "Ok Google"
   "Google"
   "I don't own this device"
   6. Do you use any of the following services on your device?
   Play music
   Set alarm and reminders
   Shopping
   Third party services
   Managing other smart home devices
   7. "Amazon Echo supports third party services called skills"?
   True
   False
   I don't own this device
   8. "Google Home supports third party apps"
   True
   False
   I don't own this device
   9. Which device has voice purchasing code?
   Amazon Echo
   Google Home
   I don't know
   10. Which device has the capability to distinguish between different speakers?
   Amazon Echo
   Google Home
   I don't know

B Interview Questions

1. Which device do you own?

2. Can you tell me about your device, what made you start using it?
   Follow-up: How long have you been using it?
   Follow -up: Other than you, who else uses it?
   3. What do you use the device for?
   Follow up: Do you use third party skills/apps?
   Follow up: What do you use?
   Follow up: How often do you use it?
   Follow up: Did you have to setup anything before you started using it?
   4. Other than your device, do you own any other smart home device?
   Follow up: Do you use your device to control your other smart home device?
   Follow up: How useful is your device in terms of controlling your smart home device?
   5. How did you register your device?
   Follow up: Was this done with your existing account? [If used an existing one]
   Follow up: Is this just for your device or you use the account for other things as well?
   Follow up: Can you tell me why you linked them?
   [If not linked]
   Follow up: Is there any reason why you didn't link them?
   [If created new account]
   Follow up: Is there a reason behind creating a new account than using an existing one?
   6. How many accounts do you have setup on your device?
   Follow up: Do these belong to others that use the device?
   Follow up: Do you use those other accounts or just one?
   [If only one account]
   Follow up: Is this shared by multiple users?
   Follow up: Can you explain why you chose to share an account?
   7. Have you completed the voice recognition process?
   Follow up: How did you find it?
   Follow up: Does the device respond to you when you speak to it?
   Follow up: When the device doesn't respond or understand you, what do you do?
   8. Can the device distinguish users or tell users apart?
   Follow up: How do you think this process works?
   Follow up: Did you experience any challenges in terms of the device identifying who you are?
   Follow up: If any, what did you do to overcome it?
   Follow up: Did you do anything to make the device recognize and identify your voice?
   Follow up: What did you do?
   9. Do you use the device to shop?
   Follow up: Can you share with me your experience in using the device to shop?
   Follow up: What do you exactly do when you shop using the device?

Amazon Echo Users only: Did you setup the voice purchasing code?

Follow up: Can you describe your experience setting up your purchasing code?

Follow up: Is your voice purchasing code always enabled?

[If disabled]

Follow up: Can you tell me why you have it disabled?

Follow up: What are your thoughts on shopping using the device?

10. Scenario 1 - Built in Services

When[NAME OF BUILT IN SERVICES] how does the device get the information you requested?

Follow up: Do you know if these requests are stored?

Follow up: [If yes] where do you think they are stored and for how long?

Follow up: Do you think Amazon or Google use this data for any purposes?

Follow up: Do you think Amazon or Google share your data with third parties like advertisers?

Scenario 2: Managing other smart home devices

You have mentioned that you use your device to manage your other smart home device, can you describe to me how you think this process works?

Follow up: Do you think what you do [activity history] are stored?

Follow up: [If yes] where do you think they are stored and for how long?

Follow up: Do you think your device shares data with [NAME OF THE OTHER DEVICE COMPANY]?

Follow up: Do you think Amazon or Google and [NAME OF THE OTHER SMART HOME DEVICE] share your data with other third parties such as advertisers?

Scenario 3: Third Party Apps You have mentioned that you use third party skills/apps on your device [NAME] can you describe to me how you think this process works?

Follow up: How does Alexa or [Google] communicate with [NAME OF App]?

Follow up: Do you know if these requests are stored and for how long?

Follow up: Do you think the device shares date with [NAME OF THE SKILL/APP]?

Follow up: Do you think Amazon or [Google] and [NAME OF THIRD PARTY SKILL/APP] share your data with other third party companies such as advertisers?

Scenario 4: Voice Purchasing

You mentioned that you sometimes use your device to purchase online, can you briefly describe to me how you think this process works?

Follow up: Do you know if purchasing orders are stored?

Follow up: [if yes] where do you think they are stored and for how long?

Follow up: Do you think Amazon or Google use this data for any purposes?

Follow up: Do you think this data is shared with other third parties like advertisers?

11. DO you think the device is able to learn things about you based on what you have asked before?

Follow up: How do you think the device is able to do that?

Follow up: Can you give me an example of what you think it has learned about you previously

Follow up: What are your thoughts on the device being able to learn things about you that you may not have said to it explicitly?

Follow up: Where do you think what the device learns about you is stored?

Follow up: Do you think Amazon or Google could use what the device learns about you for any purposes?

Follow up: Do you think what the device learns about you is shared with third parties e.g. advertisers?

12. Do you think the device could be exploited maliciously?

[If yes]

Follow up: Who do you think would be interested in exploiting the device?

Follow up: What do you think their motive is?

[If no]

Is there any specific reason you think it cannot be exploited?

13. How do you protect yourself from the device or those who might attack it?

Follow up: How effective is that?

14. Do you have any concerns on how the device handles your data?

[If any concerns]

Follow up: Does that impact the way you use the device?

15. Have you ever experienced any conflicts with others that have access to your device?

16. Have you previously experienced any incidents were the device has done something without you activating it?

17. Is there anything else you do apart from what we have talked about already?