Towards Real-time Voice Interaction Data Collection Monitoring and Ambient Light Privacy Notification for Voice-controlled Services

Tu Le*, Zixin Wang†, Danny Yuxing Huang‡, Yaxing Yao§ and Yuan Tian¶
*University of California, Irvine
†Zhejiang University
‡New York University
§Virginia Tech
¶University of California, Los Angeles

***Abstract*—Voice-controlled devices or their software component, known as voice personal assistant (VPA), offer technological advancements that improve user experience. However, they come with privacy concerns such as unintended recording of the user's private conversations. This data could potentially be stolen by adversaries or shared with third parties. Therefore, users need to be aware of these and other similar potential privacy risks presented by VPAs. In this paper, we first study how VPA users monitor their voice interaction recorded by their VPAs and their expectations via an online survey of 100 users. We find that even though users were aware of the VPAs holding recordings of them, they initially thought reviewing the recordings was unnecessary. However, they were surprised that there were unintended recordings and that they could review the recordings. When presented with what types of unintended recordings might happen, more users wanted the option to review their interaction history. This indicates the importance of data transparency. We then build a browser extension that helps users monitor their voice interaction history and notifies users of unintended conversations recorded by their voice assistants. Our tool experiments with notifications using smart light devices in addition to the traditional push notification approach. With our tool, we then interview 10 users to evaluate the usability and further understand users' perceptions of such unintended recordings. Our results show that unintended recordings could be common in the wild and there is a need for a tool to help manage the voice interaction recordings with VPAs. Smart light notification is potentially a useful mechanism that should be adopted in addition to the traditional push notification.**

I. Introduction

The Internet of Things (IoT) has increasingly made its way into our daily lives, providing lots of convenience and improvement to our quality of life. An important feature of IoT is the voice control capability, which allows the devices to "listen" to users' voice commands and execute various operations. The voice capability is usually integrated into the devices as a type of software called voice personal assistant (VPA). VPAs can significantly increase searching efficiency, quality of decision-making, and the e-commerce economy by simplifying the purchasing process [6]. Moreover, VPAs lower the bar of required technical skills to operate - one only needs to give a voice command after all. This has exciting implications for elderly and cognitively impaired individuals [37]. Although VPAs bring a lot of benefits to our lives, many privacy concerns have arisen [7], [5], [40].

Problem. The problem of consumers' unintended (if not private) conversations being recorded has become a significant privacy concern as voice assistants and voice-controlled devices are getting more popular in our world of smart technologies. IoT systems with voice control capability such as smart speakers, smart TVs, and smart home security systems have become integral parts of our lives, providing convenience and improvement to our quality of life. However, with the increasing prevalence of such systems, users are subjecting themselves to potential surveillance. There have been cases where the devices mistakenly interpret background conversations or noises as voice commands, leading to unintentional recordings. This raises serious privacy issues such as unauthorized data collection or misuse of personal information, as these recordings may contain sensitive information that users did not want to share. The lack of transparency and control over these recordings amplifies the risks since users may not even be aware of what has been recorded or how it will be used. Therefore, it is necessary to improve privacy measures and provide helpful guidelines to protect user privacy.

Despite all of these issues, many users are not aware of how easily their privacy can be compromised. They may not fully understand the extent to which these devices are listening and recording, and the potential risks associated with unintended recordings. Dubois et al. [11] revealed that even something as simple as a loud TV program can wake VPAs, causing them to record your conversations without your knowledge for up to 10 seconds. If hackers get a hold of these unwanted and private recordings through security attacks, users may be put at risk [5]. However, it is not only security attacks for which users must be aware but also the service providers collecting user information with the intention of selling it to advertisers, including personal data that users may have accidentally disclosed [27]. Often, the user is not even aware that this is happening. Therefore, it is important to improve users' awareness to protect their privacy.

Symposium on Usable Security and Privacy (USEC) 2024
26 February 2024, San Diego, CA, USA
ISBN 979-8-9894372-5-2
https://dx.doi.org/10.14722/usec.2024.23039
www.ndss-symposium.org, https://www.usablesecurity.net/USEC/

Users' awareness and perceptions are underexplored. The usage of voice-controlled devices continues to grow. However, users' awareness and perceptions of unintended conversations being recorded by the devices remain underexplored. While privacy concerns related to these devices have gained attention, little work has looked into how well users understand the extent of unintended recordings and their potential risks. Exploring users' awareness is crucial in designing effective privacy controls and notifications. It is important to conduct in-depth studies to understand how users perceive the risks, their expectations regarding privacy, and their knowledge of the data collection practices of voice-controlled devices. By getting such insights, researchers and manufacturers can develop effective solutions that protect user privacy and help users make privacy-conscious decisions.

There is a lack of effective ways to manage voice interaction recordings. Service providers like Amazon Alexa provide interfaces to present interactions recorded by the devices to users. However, such interfaces are often inadequate. While they have basic controls such as viewing information about the recorded interactions and deleting the records, the process can be cumbersome and lacks transparency. Users may not fully understand the provided information, or they may not even know about the available controls. As a result, users are left with limited control over their recorded conversations, leading to concerns about privacy and the potential for misuse or unauthorized access to sensitive information. Currently, there is a lack of effective tools to help users manage interactions recorded by voice-controlled devices. The absence of a centralized and user-friendly interface that allows users to easily review, manage, and delete their recorded data is a significant shortcoming.

Traditional push notification method might not be enough. Push notifications sent to the user's connected smartphone or other devices have been a popular approach used to notify users of privacy incidents or information that they need to pay attention to. Similarly, push notifications can serve as an alert, informing the user that an unintended conversation has been recorded by their voice assistants and voice-controlled devices. The advantage of this method is that it provides users with immediate information about potential privacy issues, allowing them to take appropriate action such as reviewing and deleting the recordings. However, there are limitations to the push notification method. Users may not always have their devices nearby or may have notifications disabled, which could result in missed alerts. Additionally, users might have multiple devices set up in different rooms. Furthermore, users may become desensitized to notifications over time, leading to a decreased sense of urgency in addressing privacy concerns.

Research Goal. In this paper, we focus on how to improve users' awareness of unintended conversations recorded by their VPAs and help them manage such recordings effectively. Awareness of the potential risks helps users make more informed decisions about the use of voice-controlled devices and take appropriate precautions to protect their privacy. This includes understanding the device's default settings, reviewing and adjusting privacy options, and being mindful of the environment in which these devices are placed. Furthermore, educating users about the potential risks of unintended recordings, such as unauthorized data access or the potential for data breaches, can help them make privacy-conscious choices and demand stronger privacy protections from service providers.

Contributions. We make the following contributions to address the mentioned problems:

- **Understanding users' awareness and perceptions:** We conduct a survey to understand users' perceptions towards voice interactions history stored by VPAs and whether the users review the records as well as their expectations.
- **Helping users manage their voice history and notification via smart lights:** We build a browser extension called VPAWatcher that automatically collects users' voice interactions recorded by their VPAs and visualizes the data. VPAWatcher notifies users of any unintended recordings happening in real time. Other than the traditional push notification method, we incorporate users' smart light devices as an option to deliver notifications. Notification through smart light devices involves visual cues, such as changing the color, pattern, or status of smart lights, to indicate the presence of unintended recordings in real-time. This method provides a more visible and ambient notification that can catch users' attention even if they are not actively using their devices that were set up for push notifications.
- **Evaluating the need for managing voice interaction recordings and testing user preferences for notifications.** We conduct an interview study with our extension to understand if such a tool to help users manage their voice interactions is necessary and the preferences users have for privacy notifications about suspicious recordings.

Key Findings. Our key findings in this paper include:

- Most users did not review their recorded voice interactions. The main reasons include not knowing how to access the recordings or thinking it is unimportant to review.
- Many users were surprised that a private conversation could be recorded by their VPAs.
- Searching for a record is a difficult task to do with the existing interface provided by Amazon Alexa.
- Unintended records are actually common. We found about 6-25% of total records are unintended for the 10 participants in our interview study.
- Most users thought it was necessary to have a tool to assist them with managing voice interaction recordings. We identified expectations and design recommendations from our user studies.
- Push notifications allow retrospective, while light notifications are more natural and attention-grabbing. Smart light devices can be placed strategically throughout our living spaces, enabling notifications to be received from any corner of the room. Light notifications are preferred for highly critical notifications.

II. Related Work

This section presents our literature review and how our work is different from previous work. The related work is presented as two themes: (1) security and privacy risks of voice devices, and (2) privacy control and notifications for voice devices.

A. Security and Privacy Risks of Voice-controlled Devices

VPA has been a very popular integration for many IoT devices and applications to facilitate voice control capability. Previous research showed various issues of VPA's speech recognition systems [20], [42], [4], [32], allowing an adversary to eavesdrop on users. Other work studied app vetting mechanisms for VPA [12], [24], [25], [8], [9], [36], showing that many published apps had bad privacy practices. Some apps were found asking users for private information [18], [22]. In our work, we focus on users' interactions with their VPAs and the issue of unintended recordings. Previous studies [11], [30] identified patterns of accidental activation of smart speakers. Adaimi et al. [2] further showed privacy leaks could even come from background sounds of intentional activation. However, these studies did not investigate users' awareness of their data being recorded by the devices, and there is still a lack of a solution to help users be aware of and control such recordings, which is the contribution of our work.

Several studies looked into the security and privacy issues of smart speakers from users' perspectives, showing that users have an incomplete understanding of how smart speakers operate [1], [17] and are concerned about their privacy [21], [39], [26], [3]. Different from these studies, we explore users' awareness of their voice data collection and how to help them manage such collection.

B. Privacy Control and Notifications for Voice-controlled Devices

Designing more comprehensive and user-friendly privacy dashboards can be an effective way to give users more control over their security and privacy when using voice personal assistants (VPAs). The importance and effectiveness of privacy dashboards were highlighted by Irion et al. [19] as one of the most feasible methods of enhancing control for users and maintaining consistency with rising standards of privacy. Creating a well-designed privacy dashboard is complex, however, requiring an understanding of user demands as well as general best design practices. Farke et al. [13] surveyed users of Google's My Activity dashboard to better understand user perceptions and reactions to a privacy dashboard. Raschke et al. [29] designed a mock-up dashboard to present a possible implementation for generalized sensitive data. However, Feth et al. [15] emphasized that privacy dashboards are not one-size-fits-all, and should be tailored to the specific domain and technology. Thus, a privacy dashboard for voice personal assistants should be designed and tested independently. Sharma et al. [31] surveyed users of Google Voice Assistant, in particular, to explore the specific needs and expectations of VPA technology and controls while also designing an algorithm to classify sensitive VPA interactions. However, these previous studies have yet to design a working tool with enhanced features and evaluate it in real world. We conduct a survey on VPA users to understand user attitudes and expectations about the privacy implications of VPAs and their associated privacy dashboards. Based on the insights, we develop a new tool to help users be aware of and control the collection of VPA interaction data in a user-friendly and privacy-sensitive manner.

Privacy notifications inform users about the data collection and usage policies of a system, product, or service. Previous work looked into different methods to deliver notifications. De Russis et al. [10] studied smartphone notifications, categorizing notification modalities based on accessibility level which contains sight, hearing, and hands level. Zeng et al. [41] designed mobile app notifications for access controls in the multi-user smart home context. Little work has been done on exploring different notification modalities. Voit et al. [35] evaluated three different notification modalities (i.e., on-object, on-environment, and on-smartphone notifications) during a cooking session, showing that on-environment notifications were perceived as the least disruptive. However, they only focused on a specific activity (i.e., cooking) and did not consider privacy notifications. Recent work also looked into privacy notification preferences in smart homes [33] and smart commercial buildings [23]. Thakkar et al. [33] surveyed users and bystanders regarding their preferences for four ways to send notifications about data practices in smart homes, i.e., visual signals (e.g., LED indicator), audio cues (e.g., voice reminder), push notifications through associated apps, and interactive web apps. These findings were based on hypothetical scenarios presented in their surveys. Our work uses experiments in which our users experience real examples, and we also provide further findings about users' reactions.

III. Survey: Perceptions of Voice Interaction Recording

To better understand the awareness and concerns that users have about voice personal assistants and their ability to record conversations, we performed a survey on 100 VPA users. The survey was designed to provide insights into the following three research questions.

- **S1**: Awareness: Do users think that VPA devices record interactions and allow users to review the interaction history?
- **S2**: Actions: How do users actually review their interaction history and do they think it is necessary to do so?
- **S3**: Expectations: What expectations do users have for reviewing their interaction history?

The goal of these questions was to motivate both the creation and the design of our Voice Interaction Extension. Qualitative and quantitative analyses were performed on the survey results to reveal user opinions and sentiments that are relevant to the development of our extension. In this section, we describe our recruitment strategy, survey design, response filtering, and results. Our study was approved by our Institutional Review Board (IRB).

A. Recruitment

We recruited 100 participants on Prolific[1] and used Qualtrics[2] to build our survey. Participants were required to be age 18 or older, live in the U.S., fluent in English, and were VPA users. To ensure data quality, we made our survey a one-time survey and available to participants with at least a 95% approval rate on Prolific. We paid each participant $1 for completing our 5-minute survey.

[1] https://www.prolific.com
[2] https://www.qualtrics.com

We filtered out invalid responses such as incomplete responses (including meaningless ones that the participant entered only white spaces into all free-text answer boxes), and responses that failed our attention checks[3]. When an invalid response was removed, the spot would be open to new participants.

B. Survey Instrument

Our survey consisted of the following three sections. The detailed questionnaire is attached in Appendix A.

- **VPA Usage**. Participants were first asked which VPA services they used and were presented with multiple options, which they can select multiple (Amazon Alexa, Google Assistant, Microsoft Cortana). Participants were then asked how long they had been using VPAs and how often they used them. Participants were also asked questions about shared VPA devices in their households, including how many people share the devices and who they may share their devices with.
- **Perceptions of Voice Interaction Recordings**. This section investigated the participants' awareness of voice and interaction history recording by VPA services and usage of interaction history and management features. First, participants were asked if they believed their VPAs recorded their voice interactions. For those who believed their VPAs recorded their voice interactions, we considered two types of interactions: *intended* and *unintended*. *Intended* interactions are what the user means to say to their VPAs, while *unintended* interactions are other random conversations that the user does not mean to say to their VPAs. In particular, participants were provided with examples of both types of interactions (shown in Appendix A). We then asked what types of interactions they thought their VPAs recorded and whether they believed the service provider let them review the history of *intended*, *unintended*, or neither type of interaction. Next, participants were asked if they reviewed the history of interactions recorded by their VPAs. Participants were also asked to explain via a free-text response why they did or did not review the history. Those who responded that they did review the history were further asked additional questions about how often they reviewed the interaction history, which platforms they used to access the interaction history, and their opinions on the process of reviewing the interaction history. Besides, for the participants who did not believe their voice interactions were recorded by their VPA devices, we explained to them that the VPA devices actually recorded their voice interactions before asking about their preferences in the next section described below.
- **Preferences for Managing Voice Interaction Recordings**. This section investigated the participants' thoughts on the importance of reviewing the voice interaction recordings and their expectations. Participants were first asked if they thought having the option to review the interaction history is necessary and what information/features they would want from an interaction history page. Next, we presented a real-world scenario where a personal conversation is recorded by a VPA (shown in Appendix A). We were interested in their reactions to the scenario. Additionally, to observe whether knowing that scenario could happen would affect their decisions, we asked the same set of questions about the need to review the interaction history and what information they would want to know again.
- **Demographic Information**. We asked for basic demographic information: gender identity, age, the highest level of education completed, and comfort level with computing technology.

C. Pretest

Pretesting via pilot studies is a common practice before deployment to identify potential issues and biases in surveys, such as priming wording or confusing questions [28]. We followed an iterative review process with pilot studies of 20 participants to receive feedback and improve our survey design accordingly until no issues arose.

We improved the wording of survey questions and made important information clearer. For example, we highlighted the context/examples presented to the participants to avoid misunderstanding. We excluded the pretest data from our final results to avoid biases.

D. Data analysis

Our data included multiple-choice (only 1 choice can be selected), multiple-response (multiple choices can be selected), Likert scale, and free-text answers. We conducted descriptive statistics to report these quantitative data. Our analysis focused on investigating the users' perceptions of voice interaction history and how the users manage their voice history. For instance, we explored whether being aware of the recording makes users think there is a need to review past interactions.

Free-text responses were independently coded by two researchers using a codebook. We coded ten random responses to construct the initial codebook and continued adding new codes throughout the coding process. To ensure the quality and inter-rater agreement, we discussed and finalized the codes as a group to resolve conflicts.

E. Results

1) Demographics: Among the 100 participants, 52.0% are male, 47.0% are female, and 1% are non-binary. Our participants skewed towards young (42.0% are between 25 and 34), highly educated (58.0% completed Bachelor's degrees or above) people. Since our recruitment focused on VPA users, our participants had a technology background and experience with computing devices (18% are experts, 68% are advanced, and 14% are intermediate). Table I presents the complete demographic information of our participants.

2) VPA usage: The majority of participants used Alexa (76 participants). 46 participants used Google Assistant and 7 participants used Cortana. Among all participants, 27 participants used multiple platforms. Our participants were mostly experienced users, which means they had been using VPAs for 1-2 years (30.0%) or more than 2 years (57.0%). Most Alexa

[3]https://researcher-help.prolific.com/hc/en-gb/articles/360009223553-Prolific-s-Attention-and-Comprehension-Check-Policy

TABLE I: Survey Participants' Demographic Information

		Participants (N=100)
Gender	Male	52
	Female	47
	Non-binary	1
Age	18-24 years old	24
	25-34 years old	42
	35-44 years old	19
	45-54 years old	8
	55-64 years old	6
	65-74 years old	0
	75 years or older	1
Education	Some high school	0
	High school graduate	12
	Some college	19
	Associate's degree (2-year college)	11
	Bachelor's degree (4-year college)	33
	Graduate degree (Masters, PhD, etc.)	25
Comfort with technology	(Expert) I can build my own computers, run my own servers, code my own apps, etc.	18
	(Advanced) I know my way around computers and mobile/IoT devices pretty well; I am the person who helps friends and family with technical problems.	68
	(Intermediate) I know how to use computers and mobile/IoT devices to perform my job and life responsibilities; I often need technical help from others.	14
	(Novice) Technology usually scares me. I only use it when I have to.	0

participants used their devices at least once a day (86.8%). Most participants (74.0%) had their VPA devices shared in their households. The people that they shared their devices with mostly included their spouse/partner, their parents, and their kids.

3) Awareness of voice interaction recordings and the ability to review them (S1): **Participants were aware that VPAs record their interactions (including the unintended ones), but they did not think the service provider let them review such unintended recordings.** 78.0% of participants thought that VPAs keep recordings of their interactions while the other 22.0% thought otherwise. For the participants who thought VPAs keep recordings, we then presented an example of an intended interaction (e.g., "Alexa, what's the weather?") and an example of an unintended interaction (e.g., a sample conversation between you and your friend). We asked two follow-up questions. First, we studied what type(s) of recordings (intended and unintended) the participants thought VPAs keep recordings of. Most of them (85.9%) thought that VPAs keep recordings of both intended and unintended interactions, while the remaining 14.1% thought that VPAs only record intended interactions. Second, we asked the participants whether they thought the service provider allowed them to review the recordings of intended and/or unintended interactions. 41% thought they could review only intended interactions. 34.6% thought they were not allowed to review anything. Only 24.4% thought they could review the history of unintended interactions. Additionally, of the 85.9% who thought that VPAs keep recordings of both intended and unintended interactions, only 28.4% thought they were allowed to review the unintended recordings. We further found that, of the 14.1% who thought that VPAs only record intended interactions, 27.3% thought they were not allowed to review anything.

4) How users actually review their interaction history and whether they think it is necessary to do so (S2): **Participants thought it was not necessary to review the interaction history or did not know they could do that.** For the participants who thought VPAs record their interactions, we explored whether they actually review the recordings and how they do that. We found that most (76.9%) did not review the recordings. We further asked our participants to explain why they reviewed or did not review the recordings via free-text responses. We found the main reasons for reviewing were privacy concerns and curiosity:

"*I wanted to make sure that wasn't anything too embarrassing being recorded as well, but mainly I wanted to know what they were collecting. (P49)*"

"*I was curious what they sounded like and what was recorded. (P60)*"

Other reasons include confirming the accuracy (16.7%), giving user feedback to service providers (11.1%), or just for fun (5.6%). Figure 1 shows the percentage of responses for each reason to review the history.

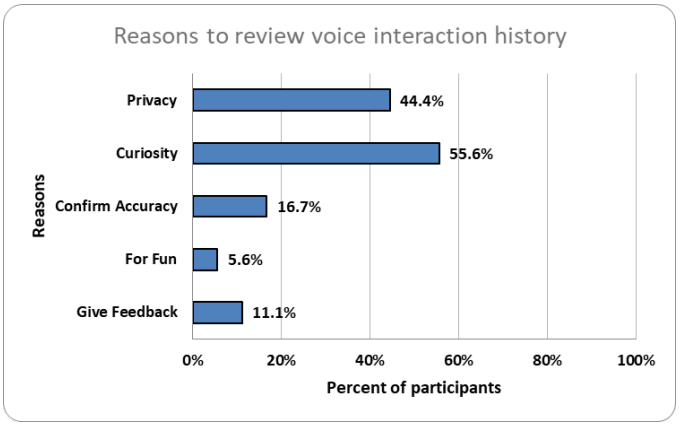


Fig. 1: Participants' reasons for reviewing their voice interaction recordings.

For why participants did not review, we found the main reasons to be the lack of awareness (i.e., did not know that they could review) (25%) and that participants thought it was not necessary (60%). Other reasons include lack of time (13.3%), thinking the interface is difficult to use (6.7%), having trust in the service provider (3.3%), and thinking it causes anxiety (1.7%). Figure 2 shows the percentage of responses for each reason not to review the history.

Participants rarely reviewed their recorded interactions. For the participants who reviewed the interactions, 82.4% reviewed once a month, and 17.6% reviewed once a week. Our participants were more familiar with the mobile app (61.1%) than the website (33.3%) when we asked how they would access the interaction history page.

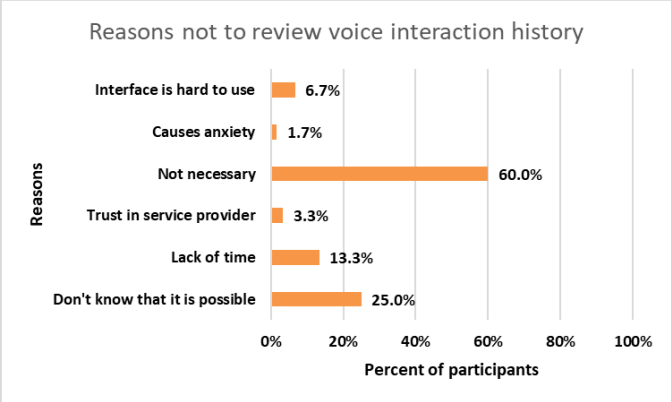


Fig. 2: Participants' reasons for not reviewing their voice interaction recordings.

Participants hardly ever deleted their recorded interactions. For the participants who reviewed the interactions, 44.4% never deleted, 33.3% sometimes deleted, 5.6% deleted about half of the time, 5.6% deleted most of the time, and 11.1% always deleted. We further asked the participants to tell us what types of interactions they would delete. Their main targets to delete were unintended interactions (including ones that may contain sensitive info) or everything:

"*Anything that was not an intended interaction, i.e., where it thought it heard its name being called and then listens in to what is being said in response. (P8)*"

"*Embarrassing queries. (P56)*"

"*Private search. (P46)*"

One participant mentioned deleting old records before a certain date:

"*I sometimes delete older recordings. (P64)*"

Participants were surprised that a private conversation could be recorded by VPAs and expressed privacy concerns. Before we presented an example of an unintended conversation recorded by an Alexa device, 83% of participants thought that it was necessary to have an option to review their voice interaction history. We explored the correlation between thinking VPAs record interactions and thinking it is necessary to have the option to review the interactions. Among the participants who thought VPAs record interactions, 88.5% also thought it was necessary to have the option to review the interactions. On the other hand, among the participants who did not think VPAs record interactions, only 63.6% thought it was necessary to have the option to review the interactions. This difference is statistically significant ($\chi^2 = 7.495, df = 1, p < 0.05$). We then explored why the other 36.4% did not find it necessary. The majority of them (82.4%) did not think it was important or beneficial; 11.8% said it would take too much time and effort to review the interactions; 5.9% were not concerned about the recordings. One participant further mentioned the lack of motivation for the service providers to provide such a feature:

"*I think companies can still successfully sell their products without needing to let us review the history. (P22)*"

After we presented an example of an unintended conversation recorded by an Alexa device, 58% of participants were surprised and concerned about it:

"*This would actually be concerning to me...this is an example I had not thought of. (P32)*"

"*I think it is wrong that it records when not spoken to. (P60)*"

"*This could pose a big threat to privacy personally and professionally if someone hacked into your recordings. (P99)*"

After seeing the example, the percentage of participants who thought it was necessary to have an option to review their voice interaction history raised from 83% to 92%.

5) Expectations (S3): **Participants found it difficult to search for a record in the interaction history.** Figure 3 shows the difficulty level that our participants selected for each task. Other than searching for a record, accessing the history interface or reviewing a record could also potentially be improved. The main problems they mentioned are that it is time-consuming and difficult to find certain records they want.

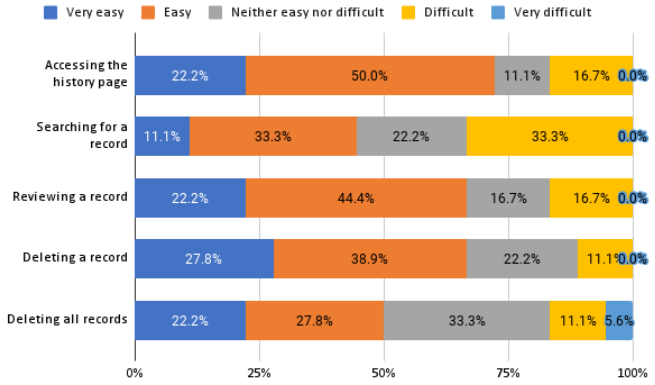


Fig. 3: Participants' opinions about how easy or difficult for them to review their voice interaction recordings.

Transparency is important. As mentioned above, the majority of participants wanted to have the option to review their voice interaction recordings. We asked the participants to report what information about the recorded interactions should be provided and/or highlighted. Figure 4 summarizes the details about the voice interaction recordings that the participants thought should be highlighted. In particular, we found that the participants preferred to know all possible info about the recordings. Some other important details about the recordings that need to be highlighted include the date, duration, what was recorded, whether the interaction was unintended or intended, transcript, sensitive info, what triggered the recording, and

if/when the recordings will be deleted. A participant also suggested highlighting common phrases to help with filtering:

> "*Maybe highlighting some of the most commonly used phrases can help users see what they say most to the assistant or can help filter these out if the user is looking for something less commonly said. (P93)*"

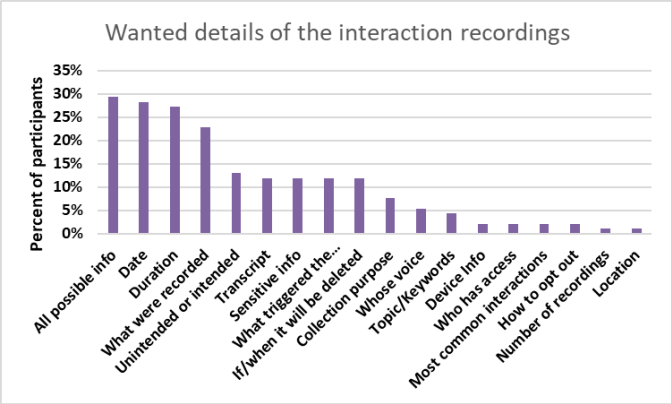


Fig. 4: Participants' suggestions for what details about the voice interaction recordings need to be highlighted.

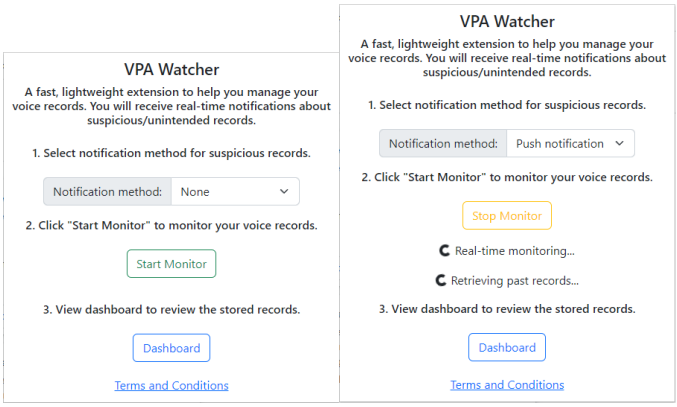
IV. Prototype: Helping Users Monitor Voice Interaction History

Our survey showed that users were unaware of the unintended interaction recordings stored by their VPAs. Although users wanted to have control over this data collection, they found it difficult to do that effectively. To get more in-depth insights into how users manage their device's data collection and their preferences, we built a browser extension to help users monitor their voice interactions with voice personal assistants. In this section, we describe the design and implementation of our VPAWatcher prototype, focusing on two dominant platforms: Amazon Alexa (voice assistant) and Philips Hue (smart light), as a proof-of-concept. VPAWatcher can be further developed to support more platforms in the future.

A. Overview

We build VPAWatcher (Figure 5) to be a browser extension that is easy to use and does not require a complex installation process. Our goal is to make the extension as simple as possible for users to easily understand how to use it. The main task for our extension is to continuously monitor the recordings of voice interactions done by VPAs. To achieve this, the extension runs in the background to identify new interactions recorded in real time. Users can also review past records.

When starting, VPAWatcher will first check if it can connect to the user's VPAs. The user will be prompted to have their VPA accounts logged in on their browser if needed. After successfully connecting to the user's VPAs, VPAWatcher is ready to serve. The features are outlined as follows:



(a) Idle (b) Monitoring

Fig. 5: Main user interface of VPAWatcher.

B. Retrieving Past Voice Interaction Records

VPAWatcher automatically sends fetch requests to collect all available voice interaction records of all VPA devices registered under the connected user account. The user can view such records in a dashboard interface that will be detailed below.

C. Setting Notification Channel

VPAWatcher employs two notification channels for the users to select: push notification and smart light notification. Users can also select "None" turn off all notifications. In our current prototype, the push notification method will pop up notifications on the current device. However, a setting can be added to allow the user to set up which device to receive push notifications in case they have multiple devices. The smart light notification method requires a one-time user authorization because VPAWatcher will control the user's smart light devices to deliver notifications. This one-time user authorization is automated by VPAWatcher to make it easy for non-experts. To initialize the process, the user first clicks a "Set up" button on VPAWatcher user interface. VPAWatcher will send a permission request to the smart light hub and inform the user to grant permission by pressing the button on the hub. Once the user presses the button on the hub, the authorization is done. A private key is generated for VPAWatcher to use to control all the connected smart lights. The user can select which smart light device to use for notifications if there are multiple devices.

D. Monitoring Voice Interactions in Real-time

VPAWatcher continuously monitors the user's voice interactions with all devices connected to the user's VPA account in real-time. By default, a fetch request is sent every second to check if there is any new interaction getting recorded. If a new interaction record is marked as unintended or "audio could not be understood", VPAWatcher will deliver a notification to the user. It is important to note that speech recognition or detecting unintended interactions is not a focus of our paper. Some studies have looked into how to detect unintended interactions

or accidental activation [11], [30]. In our study, VPAWatcher leverages Amazon Alexa's built-in detection.

E. Voice Interaction History Dashboard

The interaction records from all connected VPA devices were fetched and stored in a local IndexedDB instance. The database is updated in real-time with the monitoring feature described above. VPAWatcher includes a dashboard page to present all the records to the user (Figure 6). The dashboard provides statistics on how many total records the users have across all devices and how many of them were marked as unintended or "audio could not be understood". The user can easily search through records, show only unintended records, and filter specific records of interest using keywords, device names, timestamps, etc. There is also a delete button for each record, which essentially sends a request back to the VPA account to remove the record when clicked. However, in the interview study (Section V), VPAWatcher only does a mock deletion to preserve the participants' data.

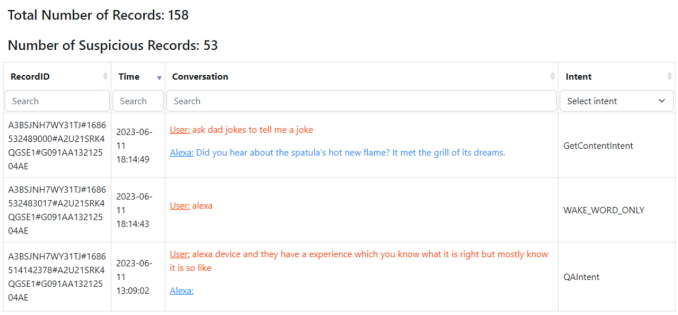


Fig. 6: Dashboard interface example showing past voice interactions with the devices.

F. Limitations

VPAWatcher is developed as a browser extension, which requires a browser to be running for it to work. However, given that browser is an essential application nowadays for daily usage, it should not be a burden on users. At the time of writing, for this study, our tool only supports Amazon Alexa and Philips Hue devices, which are in fact dominant platforms. Our tool can be developed to support other platforms in the future.

V. Interview: Voice Interaction Recording Control and Notification

To explore how people manage their voice interactions with the devices and their notification preferences, we conducted semi-structured interviews with 10 participants who are Alexa users. Since our study investigated notifications using smart lights, we recruited participants who also used smart light devices. Our study protocol was approved by our Institutional Review Board (IRB).

Our goal for the interview study is to answer the following questions.

- **I1**: How do users manage their voice history?
- **I2**: Preferences for the modality of notification (push or light) and impacts?
- **I3**: Is there a need for a tool to help manage voice history and what are the expectations?

A. Recruitment

We recruited 10 participants from Prolific. Our participants must own an Alexa device, must have a Philips Hue setup, and have experience with the devices. The participants need to be willing to use their devices and use our browser extension in the study. The participants also need to be at least 18 years old and be English speakers.

We posted our study on Prolific and used a screening survey for the participants to confirm that they met our qualification criteria. The screening takes less than 1 minute to complete (including the time to read information about our study), and the participants received $0.15 each regardless of their qualifications. We proceeded with 10 qualified participants who agreed to participate. All 10 participants completed our main study and were compensated $15 each.

Among our 10 participants, 7 were male and 3 were female. Our participants were fairly young: 2 were in 18-24 age group, 6 were in 25-34, 1 was in 35-44, and 1 was in 45-54. Most are highly educated and comfortable with technology. Our participants have been using Alexa for at least one month. Most have been using it for more than 2 years. Table II presents the demographic information of our participants.

TABLE II: Demographic information of our 10 interview participants.

	Age	Gender	Education	Comfort with technology	Alexa experience
P1	25-34	Male	Bachelor's degree	High	1-6 months
P2	18-24	Male	Associates degree	High	1-6 months
P3	25-34	Male	Bachelor's degree	Very high	2+ years
P4	35-44	Female	Graduate degree	High	2+ years
P5	25-34	Male	Graduate degree	High	2+ years
P6	18-24	Male	High school graduate	High	2+ years
P7	45-54	Female	Bachelor's degree	High	1-2 years
P8	25-34	Male	Bachelor's degree	Low	2+ years
P9	25-34	Male	Graduate degree	High	2+ years
P10	25-34	Female	Bachelor's degree	Low	2+ years

B. Design

Our semi-structured interview included several phases as follows.

1) Onboarding: We first instructed the participants to install VPAWatcher. The participants were also given a tutorial video (1 min) for reference. After the installation, we asked the participants some background questions about how they have been using their Alexa devices. These questions include how long they have been using Alexa, how often they use Alexa, what features, and if their devices are shared. Next, we asked if they thought Alexa recorded their voice interactions and how they reviewed the records. We also asked if they had any expectations for reviewing the records.

2) Experiment & open-ended discussion: The participants were asked to use VPAWatcher to check their past interaction records. Next, the participants were asked to do some interactions with their Alexa device while VPAWatcher was monitoring. The interactions include built-in functionality (e.g., "What is the weather?"), third-party skill (e.g., "Ask dad jokes

to tell me a joke"), waking Alexa up with the wake word while making some background sounds or conversations, and telling a short story with the wake word in it. Our goal was to trigger at least one unintended recording. This same task was repeated for both push notification and smart light notification mode. The participants thought out loud, let us know what actions they did, and gave us any opinions or questions they had during the entire experiment.

3) Exit questions: After the experiment, we asked some exit interview questions to evaluate the user experience and collect further comments from the participants. These questions include how difficult it is to use VPAWatcher, what they like/don't like, the need to review their interaction history, the need for a tool to assist with that, and any further comparisons of push and light notifications.

At the end of the interview, we collected basic demographic information and asked the participants to report (if they were comfortable doing it) the total number of records and the number of unintended records they have (shown on the VPAWatcher's dashboard). All but one participant reported the statistics.

C. Data Analysis

All interviews were recorded via Zoom upon participants' consent. Two researchers manually checked the transcripts to correct errors/discrepancies. We then conducted a content analysis with an open coding method guided by our research questions to identify themes and draw conclusions from the collected data.

D. Results

In this section, we detail our findings from the interviews.

1) How do users manage their voice history? (I1): All participants knew Alexa records voice interactions. However, they rarely reviewed the interactions. Most were familiar with the mobile app interface and were not aware of the website interface. Only P7 mentioned using the website interface before but it took them a lot of effort to get to it. All of them expressed some curiosity about what was recorded. However, their concerns were that there were too many entries without any filters and Alexa did not notify of any unintended records its built-in detection discovered:

> "*So many entries, there was no good way to look through the records. (P2)*"

> "*So many entries. The sensitive ones might be really buried and hard to get to. This needs some tool to help with that. (P7)*"

Most participants (all but P3) were surprised about the unintended records they had. VPAWatcher gives statistics on how many unintended records and total records the user has. Seven participants reported their number of unintended records and total records to us. Our participants had a noticeably high amount of unintended records (6-25% of total records). Table III gives the statistics on how many unintended records each participant had.

TABLE III: Seven interview participants reported their number of unintended records and total records (shown by VPAWatcher) to us. The percentage of unintended records is noticeable. Our participants have 6-25% unintended records.

	P1	P2	P4	P5	P6	P7	P8
Unintended records	120	40	2,626	725	223	3,751	2,524
Total records	1,614	370	10,411	11,098	1,394	15,836	11,179
Percentage	7.4%	10.8%	25.2%	6.5%	16.0%	23.7%	22.6%

We found that there were cases where an interaction was "secretly" recorded. This means the participants did not notice that they were being recorded due to no feedback or responses from their Alexa devices:

> "*For some of these I remember hearing random responses back from Alexa but for the others, I didn't notice the activation. (P2)*"

2) Preferences for the modality of notification (push or light) and impacts? (I2): Our participants are used to the traditional push notification method. Most (all but P4) haven't heard of notifications using smart light devices before. P4 reported that they used their smart lights to notify them of delivery or if someone is at the door:

> "*I have used light to notify us of stuff. If the lights change we know somebody is at the door. we also use it for deliveries. It's pretty helpful. (P4)*"

Participants like the idea of smart light notifications:

> "*So I find the light notifications would be way more helpful because they work all the time, even if I'm not on a computer or phone. (P4)*"

> "*Light notifications would be easier to know when it's happening. (P6)*"

Push notifications allow retrospective:

> "*In case I miss something, I can check back later because the push notification is still there. (P9)*"

However, push notifications could be easily ignored if there are multiple other notifications at the same time, while smart lights are more obvious and easy to catch attention:

> "*Light notification would get your attention because like these days like it, so many notifications on my phone. Even if it was a notification sound, it's so easy to ignore it, cause it's just another call or text or unnecessary one, right? (P7)*"

> "*I would appreciate the lights more if I wouldn't be with my push notification device at all times when a suspicious activity happens. (P2)*"

Our participants also pointed out that smart light notifications could be disruptive or annoying sometimes. Some participants thought the light notifications may disrupt their activities:

"*I haven't seen light notification in practice before. Maybe for very critical incidents, it's helpful. But If there are too many notifications or false alarms, it might be disruptive instead of informative. (P3)*"

"*The light could be disruptive in my professional duties. (P5)*"

"*Smart lights could be disruptive especially when I'm sleeping. However, it may catch my attention better if there's some serious thing happening. (P8)*"

Some participants (P3, P8, P10) preferred smart light notifications only for highly critical incidents:

"*I prefer light notification in important cases. For normal cases, the push notification is good enough. (P3)*"

A combination of both push and smart light notification methods is recommended. All participants but P5 thought it would be helpful to have both options to support each other:

"*I like the idea of a combination of different notification methods. It's flexible depending on the context like where I'm currently at when it happens. (P7)*"

P9 thought it would be helpful to set up light notifications for specific rooms:

"*I think we could have both options. Maybe light notification in other rooms if I can easily set that up would be nice. (P9)*"

This suggests that smart light devices can be set up strategically to facilitate flexible notifications.

3) Is there a need for a tool to help manage voice history and what are the expectations? (I3): It is necessary to review the recorded interactions but an automated tool would be required. After the participants used VPAWatcher, we asked them to rate the usability of our tool and whether they would be interested in using it in the future. The result shows that participants found it easy to use and most would use the tool again in the future. Figure 7 and 8 present specific details about participants' ratings for the usability of VPAWatcher.

Next, we will detail the suggestions that participants had for the tool. First, participants suggested having more notification settings. For example, flexible settings for smart light notifications (color, brightness, etc.) would be helpful:

"*Maybe provide some settings to change the color to a specific color. (P2)*"

However, this could potentially be cumbersome to some users:

"*I like the light notification but it might be confusing to set up if I have many lights. (P7)*"

"*Light notification would be helpful to notify me of serious events but I don't want to do complicated settings. It could be annoying to do the setup. (P10)*"

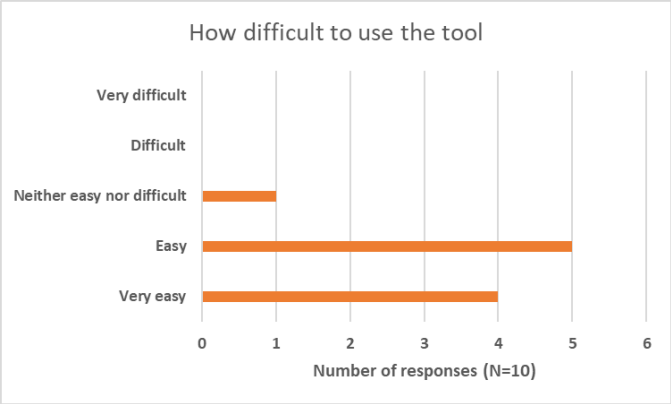


Fig. 7: Participants' ratings for how easy or difficult to use our tool.

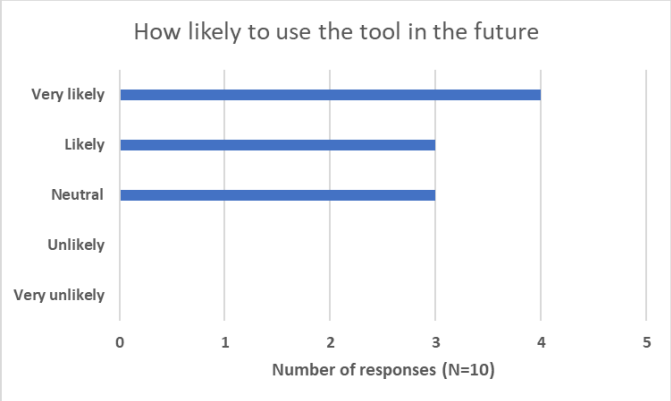


Fig. 8: Participants' ratings for how likely they will use such a tool to manage their voice interactions with the devices in the future.

Therefore, it is important to design the settings to be easy to understand and not time-consuming.

Notifications should go along with some recommended actions. Some participants were not sure about what they could do with the unintended records:

"*I don't know what to do next after seeing the records. (P3)*"

It is not transparent what controls they have with their data:

"*I resigned myself to the fact that Amazon owns my data at this point. I had no idea that I had any right to even delete this. (P5)*"

Thus, more suggestions or explanations for the users to understand their controls and the risks would help. Furthermore, P1 was clicking on the notifications hoping to get quick access to the dashboard from the notifications.

All participants would like to have automated deletions for the unintended records. The main reason was that there would be too many records to review:

"*Feels like you'd have to have some kind of service, some kind of watchdog saying you really don't need the records beyond this point because there's a lot of data to review. (P7)*"

P3 and P8 wanted to have the option to confirm before the automated deletions happen. P4 and P5 suggested having more options to delete automatically, e.g., all records from today or within a time range. Participants wanted a quick and convenient way to delete the unintended records:

"*I want 1 click to delete suspicious records automatically. (P3)*"

"*It can automatically delete unintended records for me because if it picks up every sound all day long, who wants to go through all of them? (P7)*"

P2 wanted an option to just delete everything automatically or simply a checkbox to opt out of the data collection.

VI. Discussion

This section details the implications of our paper, the ethical considerations, the limitations of our work, and our future directions.

A. Implications and Call for Action

In this paper, we studied how users perceive the recordings of their voice interactions with their VPA devices. Our findings suggest the following implications and a call for action to improve users' awareness of VPA services' data collection.

1) Voice-controlled devices' data collection is not transparent to users.: As shown in our study, although users know that the devices store data about their interactions, they are not aware of the unintended interactions being recorded. It is also unclear to users how the stored data will be used. This lack of transparency problem raises privacy/trust concerns and discomfort while around the devices:

"*I think its a bit scary because we trust so much in these companies. (P4)*"

"*It is kinda creepy that the voice assistant records all of it. I didn't even know that. (P10)*"

Therefore, it is important to improve the privacy info communication about the data collection practices and provide robust privacy controls to users.

2) A real-time monitoring tool is necessary.: Currently, there is a lack of usable tools to help VPA consumers effectively manage their interactions recorded by voice-controlled devices. The interfaces provided by service providers like Amazon Alexa lack transparency and many key features as shown in our study. Therefore, the development of a real-time monitoring tool like our VPAWatcher to manage voice data collection of devices and notify users about unintended recordings is crucial in ensuring privacy and trust. With the help of such a tool and its immediate notifications, users can be promptly made aware of potential privacy risks and take appropriate actions to mitigate the risks. Additionally, it helps improve the transparency in the use of voice-controlled devices because users can better understand when and how their interactions are recorded, which will make users more comfortable being around the devices.

3) The usability of ambient light notifications and combination of different notification modalities.: Smart light notifications offer some advantages over traditional push notifications thanks to their unique ability to integrate seamlessly into our physical world. Unlike push notifications that are confined to digital screens or audio alerts, smart light notifications provide a visual and ambient means of conveying information, which is more intuitive. Whether it is a pulsing light effect to indicate a new event or a dynamic color shift to denote a critical incident, smart lights offer a more creative way to receive notifications. We can strategically place the lights throughout our living spaces, enabling notifications to be received from anywhere we want. However, push notifications provide retrospective, which allows users to check back later if they miss a notification, e.g., when sleeping or not at home. Previous work [23], [33], [38], [34], [14] showed the need for a flexible notification strategy and that a one-size-fits-all solution would not work well in smart environments. Therefore, we envision a future notification system design that combines different modalities to support each other and allow users to have more flexible settings.

B. Ethical Considerations

We worked closely with our IRB to ensure our study protocol was in good shape. We made it clear to the participants that the participation was voluntary and that they were allowed to withdraw from the study at any time without penalty. Their responses will not be linked to their identity. We also asked our participants to freely discuss any concerns they might have about the study. Our participants did not have any concerns.

C. Limitations and Future Work

First, our user studies have some limitations due to self-reported data. We conducted several checks to mitigate the bias. In particular, we cross-checked participants' answers to ensure their responses were consistent, indicating a satisfactory level of trustworthiness regarding their opinions.

Our extension prototype in this paper only supports Alexa devices and Philips Hue devices. Thus, participants in our user studies were required to have devices from these platforms. However, this does not undermine our findings because Alexa and Philips Hue are dominant platforms. Our extension can be extended to support more platforms in the future. In addition, our tool needs permission to access the interaction history of the VPA devices and to control the smart light devices. However, our participants did not have concerns about such permissions and the setup is automated with just a few clicks. A further limitation is that the participants in our interview only interacted with a tool for a short amount of time. It is possible that using it long-term might present additional issues, such as having too many conversations to review without further guidance. Our future work can deploy the extension on a larger scale with more users and usage time, which then can facilitate some longitudinal measurements of consumers' privacy behaviors.

The data from user studies are self-reported by the participants, which means the responses might be biased due to social desirability [16]. To mitigate it, we tried to use neutral wording for our questions. We implemented attention checks to filter out inattentive participants from our survey. Furthermore, our user study protocol included incentives for completion, which might cause a bias. However, we reviewed the responses and only considered valid ones in our results. We believe that the quality of the data reported in this paper is not a problem as we tried our best to address the limitations. Besides, our user studies in this paper focus on Alexa users in the US, which is the largest user base. Our sample was from Prolific's pool of participants. Although we cannot guarantee generalizability, our findings give a lot of useful insights from the users' perspectives. We also did not investigate cultural factors. As more countries and regions adopt smart home technology, future studies can explore cross-cultural perspectives. For example, our findings regarding users' perceptions of unintended interaction recordings can be further extended to identify the differences based on different social and cultural norms.

Our research is a necessary step toward improving user awareness in the world of always-listening smart devices. Future research can look into designing personalized notification systems for privacy notices and incidents in smart environments.

VII. Conclusion

Smart home and IoT applications are becoming more popular in urban areas around the world. Such technologies often include a lot of data collection to facilitate. One popular technology is voice-controlled devices (which include Alexa-enabled voice assistant devices). These devices record a history of users' voice interactions. The recorded interactions could be either intended or unintended. This is a potential privacy concern to users but is underexplored. Therefore, our goal in this study was to understand the users' privacy perceptions of this voice data collection and their preferences for managing unintended records. Our results showed most users did not review their voice interactions and were not aware of the unintended interactions getting recorded. Users initially thought it was not necessary to review their voice interactions. However, they were surprised to see the unintended recordings and more users wanted the ability to review their interaction history. Thus, data transparency is very important. We also identified the key designs for a user interface to help with reviewing the voice interaction data and how to deliver real-time privacy notifications. Our proposed tool can help users effectively control their voice data recorded by voice-controlled IoT devices. Our findings will help guide the design and implementation of privacy-related notifications in smart homes.

Acknowledgments

This research was supported in part by the National Science Foundation (NSF) Grants (Awards 2320903, 2323105, 2317184, 2341187), the Okawa Foundation Research Grant, a Meta Research Gift, and a Google Research Gift. The US Government is authorized to reproduce and distribute reprints for Governmental purposes, notwithstanding any copyright notices thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as representing the official policies or endorsements, either expressed or implied of any funding agencies or governments. We thank all participants in our user studies for their participation and the anonymous reviewers for their comments.

References

[1] N. Abdi, K. M. Ramokapane, and J. M. Such, "More than smart speakers: security and privacy perceptions of smart home personal assistants," in *Fifteenth Symposium on Usable Privacy and Security ({SOUPS} 2019)*, 2019, pp. 451–466.

[2] R. Adaimi, H. Yong, and E. Thomaz, "Ok google, what am i doing? acoustic activity recognition bounded by conversational assistant interactions," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 1, pp. 1–24, 2021.

[3] N. M. Barbosa, J. S. Park, Y. Yao, and Y. Wang, "" what if?" predicting individual users' smart home privacy preferences and their changes." *Proc. Priv. Enhancing Technol.*, vol. 2019, no. 4, pp. 211–231, 2019.

[4] M. K. Bispham, I. Agrafiotis, and M. Goldsmith, "Nonsense attacks on google assistant and missense attacks on amazon alexa," 2019.

[5] F. Bräunlein and L. Frerichs, "Smart spies: Alexa and google home expose users to vishing and eavesdropping," Security Research Labs, December 2019. [Online]. Available: https://srlabs.de/bites/smart-spies

[6] O. Budzinski, V. Noskova, and X. Zhang, "The brave new world of digital personal assistants: Benefits and challenges from an economic perspective," *NETNOMICS: Economic Research and Electronic Networking*, vol. 20, no. 2, pp. 177–194, 2019.

[7] A. J. Campbell and L. Barrett, "In the matter of request for investigation of amazon, inc.'s echo dot kids edition for violating the children's online privacy protection act," Letter to Federal Trade Commission, Counsel for Campaign for a Commercial Free Childhood & Center for Digital Democracy, May 2019. [Online]. Available: https://www.echokidsprivacy.com

[8] L. Cheng, C. Wilson, S. Liao, J. Young, D. Dong, and H. Hu, *Dangerous Skills Got Certified: Measuring the Trustworthiness of Skill Certification in Voice Personal Assistant Platforms*. New York, NY, USA: Association for Computing Machinery, 2020, p. 1699–1716. [Online]. Available: https://doi.org/10.1145/3372297.3423339

[9] ——, "Dangerous skills got certified: Measuring the trustworthiness of skill certification in voice personal assistant platforms," in *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, 2020, pp. 1699–1716.

[10] L. De Russis and A. Monge Roffarello, "On the benefit of adding user preferences to notification delivery," in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 1561–1568. [Online]. Available: https://doi.org/10.1145/3027063.3053160

[11] D. J. Dubois, R. Kolcun, A. M. Mandalari, M. T. Paracha, D. Choffnes, and H. Haddadi, "When speakers are all ears: Characterizing misactivations of iot smart speakers," *Proceedings on Privacy Enhancing Technologies*, vol. 2020, no. 4, pp. 255–276, 2020.

[12] J. Edu, X. F. Aran, J. Such, and G. Suarez-Tangil, "Skillvet: Automated traceability analysis of amazon alexa skills," *IEEE Transactions on Dependable and Secure Computing*, 2021.

[13] F. M. Farke, D. G. Balash, M. Golla, M. Dürmuth, and A. J. Aviv, "Are privacy dashboards good for end users? evaluating user perceptions and reactions to google's my activity," in *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, Aug. 2021, pp. 483–500. [Online]. Available: https://www.usenix.org/conference/usenixsecurity21/presentation/farke

[14] Y. Feng, Y. Yao, and N. Sadeh, "A design space for privacy choices: Towards meaningful privacy control in the internet of things," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–16.

[15] D. Feth and H. Schmitt, "Requirement and quality models for privacy dashboards," in *2020 IEEE 7th International Workshop on Evolving Security & Privacy Requirements Engineering (ESPRE)*, 2020, pp. 1–6.

[16] R. J. Fisher and J. E. Katz, "Social-desirability bias and the validity of self-reported values," *Psychology & Marketing*, vol. 17, no. 2, pp. 105–120, 2000.

[17] 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*, 2018, pp. 1–6.

[18] Z. Guo, Z. Lin, P. Li, and K. Chen, "Skillexplorer: Understanding the behavior of skills in large scale," in *29th {USENIX} Security Symposium ({USENIX} Security 20)*, 2020, pp. 2649–2666.

[19] K. Irion, S. Yakovleva, J. van Hoboken, M. Thomson *et al.*, "A roadmap to enhancing user control via privacy dashboards," 2017.

[20] 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)*, 2018, pp. 33–47.

[21] 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*, vol. 2, no. CSCW, pp. 1–31, 2018.

[22] T. Le, D. Y. Huang, N. Apthorpe, and Y. Tian, "Skillbot: Identifying risky content for children in alexa skills," *ACM Trans. Internet Technol.*, vol. 22, no. 3, Jul 2022. [Online]. Available: https://doi.org/10.1145/3539609

[23] T. Le, A. Wang, Y. Yao, Y. Feng, A. Heydarian, N. Sadeh, and Y. Tian, "Exploring smart commercial building occupants' perceptions and notification preferences of internet of things data collection in the united states," in *2023 IEEE 8th European Symposium on Security and Privacy (EuroS&P)*, 2023, pp. 1030–1046.

[24] C. Lentzsch, S. J. Shah, B. Andow, M. Degeling, A. Das, and W. Enck, "Hey alexa, is this skill safe?: Taking a closer look at the alexa skill ecosystem," in *28th Annual Network and Distributed System Security Symposium (NDSS 2021). The Internet Society*, 2021.

[25] S. Liao, C. Wilson, L. Cheng, H. Hu, and H. Deng, "Measuring the effectiveness of privacy policies for voice assistant applications," in *Annual Computer Security Applications Conference*, ser. ACSAC '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 856–869. [Online]. Available: https://doi.org/10.1145/3427228.3427250

[26] H. R. Lipford, M. Tabassum, P. Bahirat, Y. Yao, and B. P. Knijnenburg, "Privacy and the internet of things," *Modern Socio-Technical Perspectives on Privacy*, p. 233, 2022.

[27] A. McCarthy, B. R. Gaster, and P. Legg, "Shouting through letterboxes: A study on attack susceptibility of voice assistants," in *2020 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*. IEEE, 2020, pp. 1–8.

[28] S. Presser, M. P. Couper, J. T. Lessler, E. Martin, J. Martin, J. M. Rothgeb, and E. Singer, "Methods for testing and evaluating survey questions," *Methods for testing and evaluating survey questionnaires*, pp. 1–22, 2004.

[29] P. Raschke, A. Küpper, O. Drozd, and S. Kirrane, "Designing a gdpr-compliant and usable privacy dashboard," in *IFIP international summer school on privacy and identity management*. Springer, 2017, pp. 221–236.

[30] L. Schönherr, M. Golla, T. Eisenhofer, J. Wiele, D. Kolossa, and T. Holz, "Unacceptable, where is my privacy? exploring accidental triggers of smart speakers," 2020.

[31] V. Sharma and M. Mondal, "Understanding and improving usability of data dashboards for simplified privacy control of voice assistant data," in *31st USENIX Security Symposium (USENIX Security 22)*. Boston, MA: USENIX Association, Aug. 2022, pp. 3379–3395. [Online]. Available: https://www.usenix.org/conference/usenixsecurity22/presentation/sharma-vandit

[32] T. Sugawara, B. Cyr, S. Rampazzi, D. Genkin, and K. Fu, "Light commands: laser-based audio injection attacks on voice-controllable systems," in *29th {USENIX} Security Symposium ({USENIX} Security 20)*, 2020, pp. 2631–2648.

[33] P. K. Thakkar, S. He, S. Xu, D. Y. Huang, and Y. Yao, ""it would probably turn into a social faux-pas": Users' and bystanders' preferences of privacy awareness mechanisms in smart homes," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, ser. CHI '22. New York, NY, USA: Association for Computing Machinery, 2022. [Online]. Available: https://doi.org/10.1145/3491102.3502137

[34] J. Vitak, M. Zimmer, A. Lenhart, S. Park, R. Y. Wong, and Y. Yao, "Designing for data awareness: addressing privacy and security concerns about "smart" technologies," in *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing*, 2021, pp. 364–367.

[35] A. Voit, T. Kosch, H. Weingartner, and P. W. Woźniak, "The attention kitchen: Comparing modalities for smart home notifications in a cooking scenario," in *Proceedings of the 20th International Conference on Mobile and Ubiquitous Multimedia*, 2021, pp. 90–97.

[36] D. Wang, K. Chen, and W. Wang, "Demystifying the vetting process of voice-controlled skills on markets," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 3, pp. 1–28, 2021.

[37] R. Yaghoubzadeh, M. Kramer, K. Pitsch, and S. Kopp, "Virtual agents as daily assistants for elderly or cognitively impaired people," in *International workshop on intelligent virtual agents*. Springer, 2013, pp. 79–91.

[38] Y. Yao, "Designing for better privacy awareness in smart homes," in *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*, 2019, pp. 98–101.

[39] Y. Yao, J. R. Basdeo, S. Kaushik, and Y. Wang, "Defending my castle: A co-design study of privacy mechanisms for smart homes," in *Proceedings of the 2019 chi conference on human factors in computing systems*, 2019, pp. 1–12.

[40] Y. Yao, J. R. Basdeo, O. R. Mcdonough, and Y. Wang, "Privacy perceptions and designs of bystanders in smart homes," *Proceedings of the ACM on Human-Computer Interaction*, vol. 3, no. CSCW, pp. 1–24, 2019.

[41] E. Zeng and F. Roesner, "Understanding and improving security and privacy in Multi-User smart homes: A design exploration and In-Home user study," in *28th USENIX Security Symposium (USENIX Security 19)*. Santa Clara, CA: USENIX Association, Aug. 2019, pp. 159–176. [Online]. Available: https://www.usenix.org/conference/usenixsecurity19/presentation/zeng

[42] N. Zhang, X. Mi, X. Feng, X. Wang, Y. Tian, and F. Qian, "Dangerous skills: Understanding and mitigating security risks of voice-controlled third-party functions on virtual personal assistant systems," in *2019 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2019, pp. 1381–1396.

APPENDIX

A. Usage

1. Which type of voice assistants do you use? (Choose all that apply)

- □ Amazon Alexa
- □ Google Assistant
- □ Microsoft Cortana
- □ None of the above

2. How long have you been using voice assistants?

- □ Less than a month
- □ 1-6 months
- □ 7-12 months
- □ 1-2 years
- □ More than 2 years

The following question was asked for each of the voice assistants below.

- Amazon Alexa
- Google Assistant
- Microsoft Cortana

3. How often do you use the following voice assistants?

- □ Never

- □ Once a month
- □ Once a week
- □ Once a day
- □ 2-10 times a day
- □ More than 10 times a day

4. Do you own voice assistant device(s) in your home that are shared among multiple people?
- □ Yes
- □ No

5. How many people use your voice assistant device(s) in your home?
- □ 1 (Only I use my device(s))
- □ 2
- □ 3
- □ 4
- □ 5 or more

6. Please indicate people that you share your voice assistant device(s) in your home with. (Choose all that apply)
- □ My parent(s)
- □ My grandparent(s)
- □ My spouse/partner
- □ My kid(s) - aged 1 to 13
- □ My kid(s) - aged 14 to 18
- □ My sibling(s)
- □ My relative(s)
- □ My guests (e.g., friends, visitors)
- □ My housemate(s) or roommate(s)
- □ Other (please specify) _____
- □ None. Only I use my voice assistant device(s)

B. Perception and Experience

7. Do you think that voice assistants keep recordings of your interactions?
- □ Yes
- □ No

If "No" to Q7, skip to the end of subsection.

Participants were then shown the following examples:

- Example of an **intended interaction**: You say, "Alexa, what's the weather", and Alexa responds with the weather info.
- Example of an **unintended interaction**: You and your friend are talking to each other about "going out for dinner this weekend", and the nearby Alexa responds with some restaurant suggestions.

8. How do you think the voice assistants keep recordings of your interactions?
- □ The voice assistants keep recordings of **BOTH** the interactions that are **intended and unintended** for them.
- □ The voice assistants keep recordings of **ONLY** the interactions that are **intended** for them.

9. Which of the following do you think is available for voice assistant users? (Choose all that apply)
- □ The voice assistant service provider **allows** me to review the history of all **intended** interactions with my voice assistant.
- □ The voice assistant service provider **allows** me to review the history of all **unintended** interactions with my voice assistant.
- □ The voice assistant service provider **does not allow** me to review the history of any of my interactions with my voice assistant.

10. Do you review the history of interactions recorded by your voice assistants?
- □ Yes
- □ No

If "Yes" to Q10:

11. Please briefly explain the reasons why you review the history of interactions recorded by your voice assistants.
_____

If "Yes" to Q10:

12. How often do you review the history of interactions recorded by your voice assistants?
- □ Never
- □ Once a month
- □ Once a week
- □ Once a day
- □ 2-10 times a day
- □ More than 10 times a day

If "Yes" to Q10:

13. How do you review the history of interactions recorded by your voice assistants? (Choose all that apply)
- □ Website
- □ Mobile App
- □ Other: _____

If "Yes" to Q10:

14. How often do you delete interaction records in the history of interactions recorded by your voice assistants?
- □ Never
- □ Sometimes
- □ About half the time
- □ Most of the time
- □ Always

If "Yes" to Q10 and not "Never" to Q14:

15. What kinds of interaction records do you delete?
_____

If "Yes" to Q10:

16. How often do you delete ALL interaction records in the history of interactions recorded by your voice assistants?
- □ Never
- □ Sometimes

□ About half the time
□ Most of the time
□ Always

The following question was asked for each of the actions below.

- Accessing the history page
- Searching for a record
- Reviewing a record
- Deleting a record
- Deleting all records

If "Yes" to Q10:

17. What do you think about the process of reviewing the history of interactions recorded by your voice assistants?
□ Very easy
□ Easy
□ Neither easy nor difficult
□ Difficult
□ Very difficult

18. Any other comments you have about the process of reviewing the history of interactions recorded by your voice assistants?
_____

If "No" to Q10:

19. Please briefly explain the reasons why you do not review the history of interactions recorded by your voice assistants.
_____

C. Preference

Participants who answered "No" to Q10 were shown the following notice before further questions were asked.

- In fact, the voice assistants actually keep recordings of your interactions. Please answer the following questions.

20. Do you think it is necessary to have an option to review the history of interactions recorded by your voice assistants?
□ Yes
□ No

If Yes to Q20, ask:

21. What information about the recorded interactions do you think should be provided and/or highlighted in the interaction history page?
_____

If No to Q20, ask:

22. Please briefly explain the reasons why you think it is not necessary to have an option to review the history of interactions recorded by your voice assistants.
_____

Participants were then shown the following example:

- **Here is a real example from the experience of a user with his Alexa device. Please read it carefully and answer the follow-up questions.**
- The user had his Alexa device in his office. One day he was curious and checked a few recent interaction records in the interaction history page of his Amazon account. He found that there were some records of what he spoke to his colleagues in an online meeting. These records included audio recordings that he could play back.

23. Given the example, do you think it is necessary to have an option to review the history of interactions recorded by your voice assistants?
□ Yes
□ No

If Yes to Q23, ask:

24. What information about the recorded interactions do you think should be provided and/or highlighted in the interaction history page?
_____

If No to Q23, ask:

25. Please briefly explain the reasons why you think it is not necessary to have an option to review the history of interactions recorded by your voice assistants.
_____

26. Other thoughts you have about the example?
_____

D. Demographics

27. Which gender identity do you most identify with?
□ Male
□ Female
□ Other: _____
□ Prefer not to answer

28. What is your age?
□ 18-24 years old
□ 25-34 years old
□ 35-44 years old
□ 45-54 years old
□ 55-64 years old
□ 65-74 years old
□ 75 years or older
□ Prefer not to answer

29. What is the highest level of education you have completed?
□ Some high school
□ High school graduate
□ Some college
□ Associate's degree (2-year college)
□ Bachelor's degree (4-year college)
□ Graduate degree (Masters, PhD, MD, JD, etc.)
□ Other: _____
□ Prefer not to answer

30. Please select the statement that best describes your comfort level with computing technology.

- □ I can build my own computers, run my own servers, code my own apps, etc.
- □ I know my way around computers and mobile/IoT devices pretty well; I am the person who helps friends and family with technical problems.
- □ I know how to use computers and mobile/IoT devices to perform my job and life responsibilities; I often need technical help from others.
- □ Technology usually scares me. I only use it when I have to.