Privacy Analysis of Voice User Interfaces

Farida Yeasmin
Tampere University
Espoo, Finland
farida.yeasmin@tuni.fi

Sneha Das, Tom Bäckström
Aalto University
Espoo, Finland
sneha.das, tom.backstrom@aalto.fi

Abstract—**While the popularity of voice user interfaces (VUIs) is increasing steadily, there is a lack of understanding about their impact on privacy. This has resulted in the rise of privacy concerns among users of VUI. Such privacy concerns include unwanted location tracking, unwarranted fingerprinting of voice data, listening to the users' private conversations without consent, and unwanted sharing of private data with other devices and services. In this paper, we present user research on the emotional experiences and privacy expectations of users from human-to-VUI interaction. Further, we investigate user preferences for privacy notification modalities in a VUI, with respect to different privacy contexts. We use a sequence of qualitative and quantitative data analysis techniques to identify these aspects from our user study. To validate our findings, we implement a prototype of a privacy-aware notification system for a VUI based application and evaluate its effectiveness. Finally, we provide guidance on improving privacy awareness of a VUI device.**

I. Introduction

A voice user interface (VUI) is a collection technologies that enables humans to communicate with a computer using their voice and speech, in order to initiate services or processes. We can use VUIs for several purposes such as controlling smart home devices, finding information, and interacting with others. A device that primarily uses a VUI for interaction with users is called a voice assistant. Recently, voice assistants have gained popularity largely due to an intuitive user interface. However, this popularity together with a lack of understanding on the private-data handling practices of a voice assistant raises privacy concerns for users. For example, a voice assistant can track and gather geographical location, habits and preferences of a user who interacts with it. The private information could then be used by companies to gain unethical advantage over people, for instance, heightened prices and denial of services targeted towards certain groups of individuals.

Traditionally, VUIs or more specifically voice assistants follow a privacy policy and notify users about privacy events based on that policy. However, for a user, in many instances it is difficult to understand written privacy policies. Therefore, written privacy policies are often ineffective in conveying the privacy practices of a VUI [1], [2]. Additionally, users mostly ignore out of context privacy policies and notifications. To address these issues, summary notices have been proposed as an alternative to long, unreadable privacy policies [3], [4]. Furthermore, context aware privacy policies and notifications have been proposed to adapt to the privacy expectations of the user based on context [5]. For instance, a user can expect that while a gaming application would not collect medical data, a health monitoring application can collect such data.

A user of a VUI may assume that interacting via VUI is similar to human-to-human communication as they both primarily use voice for interaction. Consequently, users may have false expectations and assumptions about privacy when interacting with a VUI, i.e., they may expect that human-to-VUI and human-to-human communication provide similar types of privacy. Moreover, users may have different privacy expectations with regard to *different types* of private data. However, VUIs may not follow the user's privacy expectations. For example, a user may expect that the location data of the user can be public with consent while conversation data should always remain private. However, VUI devices may record and share the user conversation with other devices and services. For instance, users' conversations may be uploaded to a cloud server and analyzed by real humans [6]. The user may be unaware that their private conversations are being recorded and analyzed by other individuals. Similarly, users may misunderstand privacy polices of VUIs, e.g., policies regarding how data is *collected* and *used* [7]. These types of mismatches between the expectation of a user and the actual privacy practices of a VUI can result in privacy breaches.

It has been reported that VUIs and their service providers have performed actions without proper understanding of the user commands [8]. This increases privacy concerns for VUIs as they can breach users' privacy by using their data without a meaningful consent. Consecutively, research in the privacy expectations of users to such scenarios as well as user's emotional response with regard to privacy when using a VUI is therefore in high demand. Note that such expectations also include, in addition to what data service providers collect and how it is used, users expectations about how matters of privacy are communicated to the user.

Privacy concerns related to voice assistants has been studied in a limited number of prior publications. A recent study on the privacy concerns associated to the use of voice activated personal assistants in public spaces was presented by Moorthy et al. [9]. It was found that people are more concerned when they transmit private information from a public place, and prefer to transmit information in a private location than in public locations. However, this study was limited in context and granularity of privacy. Additionally, the study did not focus on the users emotional experiences and expectations on privacy-aware notification systems in the context of transmitting private information over a voice assistant.

Through this work, we aim to develop an understanding of privacy in the perspective of VUIs. Therefore, in this paper, which is derived from the work presented in [10] we investigate the following questions:

1) What are the emotional experiences and privacy expecta-

TABLE I. SEVEN TYPES OF PRIVACY AS PROPOSED IN [14]

Privacy types	Examples
Privacy of the person	Voice identification
Privacy of behavior and action	Habit
Privacy of data and image	Listening to a private conversation
Privacy of thoughts and feelings	Emotions
Privacy of location and space	Location tracking
Privacy of communication	Interception of wireless communication
Privacy of association	Member of a group

tions of the user, when communicating using a VUI?

2) How do the above experiences and preferences vary for changing contexts?
3) What are the suitable modalities for privacy notifications in a VUI and do these vary with changes in context?

II. METHODOLOGY

A. Research questions

At a personal level, privacy is a sensory, subjective and contextual concept, and depends on aspects such as feelings, preferences, and deeds. Under a given situation, these aspects determine our subjective preference of privacy. Since human emotions are so critical in determining privacy, it has been part of the discussion in recent works on privacy [11], [12]. Additionally, a user has a set of privacy expectations when communicating with a VUI; for instance, a user may expect that an online service in the VUI will not store private data without the explicit consent of the user, or the data will not be shared with other online services. Furthermore, contextual differences such as location, individual traits, and groups can impact privacy expectations e.g., the privacy expectation of a user in a home environment can be different from the privacy expectations in a public space. Lastly, state-of-the-art systems notify the users on privacy practices of a service through written privacy policies. Lengthy privacy policies are time consuming and difficult to read, and are often ignored by the users [1], [2]. Past research proposes summary privacy notices in a visual format as an alternative to lengthy privacy policy statements [3]. However, VUI devices often lack graphical user interfaces (personal assistants), whereby conventional text and image based privacy notices are not directly applicable. Therefore, there is a need for privacy notification modalities which can be applied to VUI devices. With this motivation, the first part of our research addresses the identification of emotional experiences and privacy expectations of a user when interacting with a VUI. In the second part, we study the impact of contexts on the privacy experiences and expectations. Finally, we investigate suitable notification modalities for VUI devices.

B. Privacy types and contexts

Past research shows that people are more concerned about the unauthorized collection, retention, and sharing of personal data [13]. In another research, the authors classify privacy into seven categories based on the data type [14]; the categories are presented in Table I. For our study, we select the privacy types which are most relevant for human-to-VUI interaction as listed here: 1) privacy of the person, 2) privacy of behavior and action, 3) privacy of data and image, 4) Privacy of thoughts and feelings, 5) Privacy of location and space.

TABLE II. SCENARIOS BASED ON PRIVACY TYPES, CONTEXTS AND VISUAL ORIENTATION

Examples of privacy types (Tab. I)	Context	Visual orientation
1) Voice identification 2) Habits 3) Location and space 4) Private conversation 5) Emotions	Home alone and own device	Looking or Not looking
1) Voice identification 2) Habits 3) Location and space 4) Private conversation 5) Emotions	Classmate's house and guest device	Looking or Not looking

For human-to-VUI communication, we define two high-level contexts based on the location of the user. In the first context, user is at own home and communicating with the voice assistant. In the second context, user is at a classmate's house and communicating with the guest device. Furthermore, in order to use the privacy experiences from the human-to-human scenario as a reference for human-to-VUI communication, we define two comparable contexts for human-to-human communication. These are 1) the user is interacting with a close relative and 2) the user is interacting with a friend of a classmate. Within the framework of human-to-human communication, we can assume that a relative is analogous to a voice assistant located a user's own home. Similarly, a friend of the classmate can be considered as a voice assistant that is located at friends home. Note that the responses to the human-to-human communication questionnaire is not presented in this paper due to space constraints, and can be accessed in [10].

To investigate suitable modalities for privacy aware notification systems in VUI, we use the contexts for human-to-VUI interaction as mentioned above. In addition, we further increase the granularity of the contexts by incorporating two visual orientation modes, which are (1) user looking at the voice assistant, and (2) user not looking at the voice assistant, while communicating with the VUI. We assess the user responses for suitable notification modalities based on these contexts. Table II shows the scenarios based on the privacy types and contexts for identifying users privacy expectations and suitable privacy aware notification modalities while communicating with a VUI.

C. User study

In this work, we perform a user research that consists of two user studies; the first study was conducted in the exploratory phase to investigate our research questions and the second study evaluated the effectiveness of the notification system prototype. The design of the user study in both the phases was motivated by the Nielsen Norman model [15]. We interviewed eight participants in the exploratory phase and five participants in the prototype evaluation phase. The interviews were performed in a semi-structured format within a face-to-face lab setting. We collected both qualitative and quantitative data from the user study and use them for the data analysis [16], [17]. We perform both the qualitative and the quantitative analysis in the exploration phase while the prototype evaluation phase uses only quantitative analysis. The qualitative analysis uses thematic approach [18] while the quantitative analysis uses cross-tabulation approach to analyze the data. In the second study, we evaluate the effectiveness of the prototype based on the psychometric [19] response of a user.

III. Result and Discussion

A. *Emotional experiences and privacy expectations of VUI*

In this section, we present the results of user study on the emotional experiences and privacy expectations in human-to-VUI interaction for five types of privacy, and compare responses between the two human-to-VUI contexts described in Sec. II-B. Due to space constraints, we only depict the responses for location tracking and listening to private conversations in Tables. III & IV; responses for the remaining privacy types are provided in [10]. Table V summarizes the overall emotional experiences of human-to-VUI interaction for the five privacy types.

1) Privacy of the person voice identification: When an unknown voice assistant (e.g., classmate's device) performs voice identification, it leads to an emotional experience of fear amongst most participants. In contrast, participants show a positive emotion of trust, when their own voice assistant performs voice identification. As most of the participants trust their own voice assistant, the privacy expectations from the voice assistant is that their own voice assistant should be able to recognize their voice. Additionally, classmates voice assistant also can recognize their voice only if they have used the classmate's voice assistant several times. However, a participant does not expect an unknown, unused voice assistant to identify the voice. This implies that participants do not like voice identification data being shared between multiple voice assistants. Furthermore, participants expect that the voice assistant should have functionalities to remove the identified voice data and the voice assistant should take explicit consent before performing identification tasks.

2) Privacy of behavior and action habits detection: Most participants show fear when a classmates voice assistant performs habit detection. However, some participants seemed impressed with the idea of a habit detection algorithm, due to its sheer complexity. On average, participants showed negative emotions towards the idea of classmates voice assistant performing habit detection. In contrast, participants feel relaxed if their own voice assistant knows about their ordering habits and mostly indicated positive emotions for the same. In terms of the privacy expectations for habit detection, participants mostly show curiosity about the functioning of the algorithm and data collection practices. In addition, participants expect that the device should not share the data relevant to habit detection with other devices, and an option to disable the feature should be available.

3) Privacy of location and space location tracking: For both the contexts, i.e., location tracking through classmates voice assistant or own voice assistant, participants showed either distrust or anger. They stated that location tracking through a voice assistant is a serious breach of privacy. As for the privacy expectations, participants support the requirement of appropriate consent before the devices perform location tracking. This privacy expectation is consistent for both contexts. We present the emotional experience themes, categories, and privacy expectations for location tracking in Table III for eight participants.

TABLE III. Emotional experiences and privacy expectations for location tracking (N = 8)

Category	Emotional experience themes and Privacy expectations	Human-to-VUI	
		Classmates VUI (Count)	Own VUI (Count)
Distrust	Security risk	3	
	Violation of the privacy		1
	Data can not be shared		3
Sad	Unhappy	1	1
Angry	Stop using the voice assistant	4	1
	Enraged	3	2
	Not acceptable	1	3
Expectation	Require consent for data use	1	4

4) Privacy of data and image listening to private conversation: Participants react with distrust and anger to the idea of a voice assistant listening to private conversations, and stated that storing and listening to the private conversation is not acceptable in either cases. As for the privacy expectations, users should possess control over when the voice assistant listens to conversations and that should be based on explicit user consent. Additionally, private conversations should not be shared with other devices or services, and responsible data handling practices is expected from a VUI device. The responses from the user study are depicted in Table IV.

TABLE IV. Emotional experiences and privacy expectations for listening to private conversations (N = 8)

Category	Emotional experience themes and Privacy expectations	Human-to-VUI	
		Classmates VUI (Count)	Own VUI (Count)
Fear	Shocking	1	
	Fear of storing data	1	1
	Security risk	1	1
Distrust	Violation of the privacy	1	
	Disturbing	3	
Angry	Turn off the voice assistant	2	1
	Enraged	4	1
	Not acceptable	3	2
Surprise	Confusing		1
Expectation	Require consent for data use	2	
	Sharing insensitive privacy data is ok	1	

5) Privacy of thoughts and feelings emotion detection: Participants show mixed responses for the scenario of emotion detection in a VUI device, for both own and classmate's device. In other words, some participants believe it is a positive action, while other participants state it is a breach of privacy. Finally, participants expect that the collected data should only be utilized for positive purposes and the data should not be shared with other voice assistants.

TABLE V. Summary of emotional experiences

Privacy category	Human-to-VUI	
	Classmates VUI	Own VUI
Voice Identification	Fear	Trust
Habits detection	Fear	Trust
Location tracking	Angry	Angry
Listening to private conversation	Angry	Angry
Emotion detection	inconclusive	inconclusive

B. *Suitable privacy notifications*

In this section, we discuss the user responses for the preferred privacy notification modalities in the human-to-VUI framework. The responses are depicted via a radial stacked barplot in Fig. 1 over each privacy type, and for the following permutations of context and visual orientation: 1) own house, looking at the voice assistant, 2) own house, not looking at the voice assistant, 3) classmate's house, looking at the voice assistant, 4) classmate's house, not looking at the voice assistant. Each bar in the stacked barplot represents the total preference for a single notification modality, and the preference for the modality for each privacy type. An analysis of the preferences is described below.

1) Own house, looking at voice assistant: For emotion detection, listening to private conversations, and voice detection, participants choose visual notification modalities. In addition, for location tracking and habit detection, notification via app and audio were most preferred. On average, participants mostly preferred visual notification modality.

2) Own house, not looking at voice assistant: Under this scenario, participants showed a preference for app-based notification, specifically for habit detection and location tracking. For emotion detection and private conversations, participants voted for visual and audio notification modalities, respectively. Overall, visual, audio and app-based notifications are most preferred by the participants. Furthermore, we can conclude that the preference for audio notification, which is also the most intrusive notification modality, for listening to private conversations correlates to the perceived sensitivity of the privacy type.

3) Classmates house, looking at voice assistant: In the context of interacting with the voice assistant at classmates house while looking at it, participants mostly prefer audio notification for voice identification, listening to a private conversation, and emotion detection. Again, application based notification is preferred for location tracking and habit detection.

4) Classmates house, not looking at voice assistant: App-based notification is preferred for location tracking and listening to a private conversation. Most participants seem to prefer audio notifications for voice identification and emotion detection. On average, app-based notification is found to be most preferred and is closely followed by audio notification.

C. Notification system prototype

The exploratory user study presented in the former sections indicate that for classmate's voice assistant, participants prefer audio-based notifications. Thus, we developed an audio-based notification and confirmation dialog, which notifies users before storing private data. The workflow of the application is illustrated in Fig. 2 and the exact implementation details are provided in [10]. The application works on Alexa powered VUI devices and is developed on the principles of privacy-by-design [7].

1) Evaluation of the prototype: To evaluate the prototype, we perform a user study with five participants, in order to validate the effectiveness of the prototype. The prototype was designed to provide: 1) an explicit audio-based notification when private data is stored or the user granted permission to store private data, and 2) distinct audio sound (beep) to notify users that the voice assistant has performed a privacy-sensitive operation. Each notification is evaluated in the following four dimensions: 1) helpful, 2) noticeable, 3) required, and

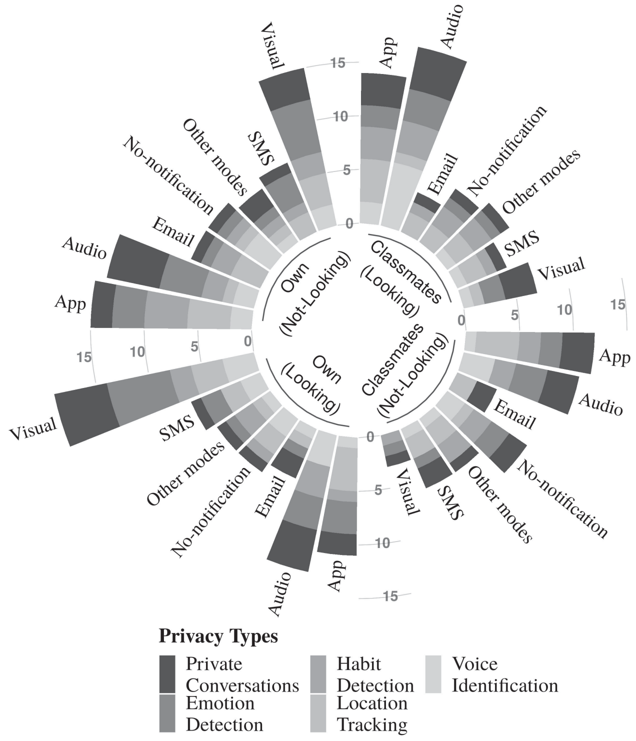

Fig. 1. Responses from the study on user preferences for privacy notification modalities in a VUI device

4) sufficient. A 5-point Likert-scale ranging over 1) Totally disagree, 2) Disagree, 3) Neutral 4) Agree, and 5) Totally agree is used to evaluate the four dimensions. From the responses shown in Table. VI, most participants *totally agree* that both the notifications are helpful and noticeable when the voice assistant stores private data from the user. All participants agree that both notifications are required to identify the privacy actions by the voice assistant. Additionally, participants believe that the audio confirmation and the beep audio are sufficient to notify users about the privacy sensitive actions performed by the voice assistant.

TABLE VI. RESULTS FROM PROTOTYPE EVALUATION

Qualitative metric to evaluate prototype	Participant					Median
	#1	#2	#3	#4	#5	
Audio confirmation is helpful	5	5	4	4	5	5
Beep audio is helpful	5	4	4	4	5	4
Audio confirmation is clear to understand	5	4	4	5	5	5
Beep audio is noticeable	4	5	5	4	5	5
Audio confirmation and beep audio is required	5	2	5	4	4	4
Audio confirmation and beep audio is sufficient	5	4	4	2	4	4

In addition to responses, participants also provided suggestions and comments to improve the notification system: 1) Participants expect that if the voice assistant stores private data, it can be checked later on. One participant stated that, "store it in the application, so it could be later checked." 2) The beep audio should happen before asking to store the private data, such that the user is aware of the impending privacy-sensitive action beforehand. In that way, the user would be more focused on what the voice assistant asks. 3) Users also showed concern about security risks, i.e, notifications should specify if the stored personal data is only used by the user's voice

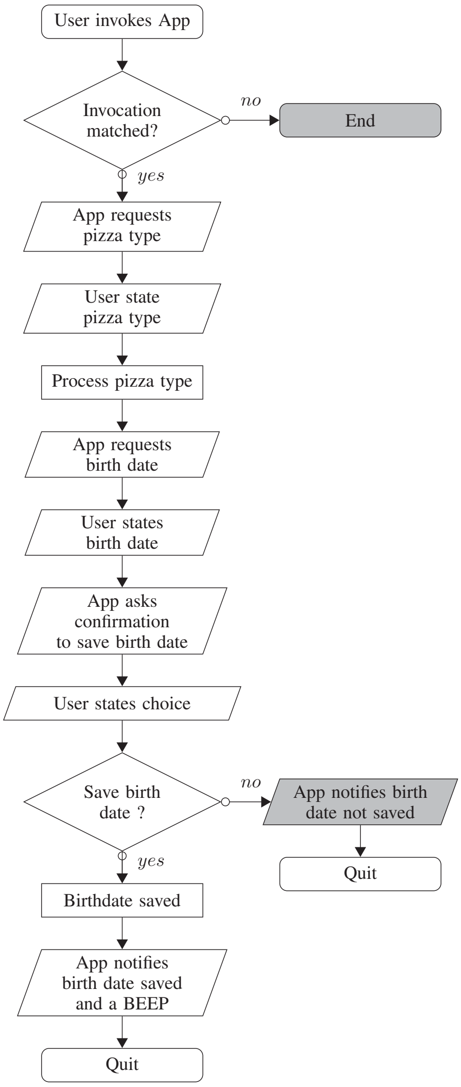

Fig. 2. Flow chart of the prototype for audio-based privacy aware notification system, implemented in an Alexa powered VUI device

assistant for the intended service or by other services as well.

D. Summary

The first user study showed that users are most concerned about privacy breaches concerning location tracking and private conversations. Furthermore, we learned the main privacy expectations of users for human-to-VUI interaction: 1) Require consent from the user when using private data. 2) Feature to forget the collected private data. (3) Feature to turn off the voice assistant. (4) Private data collected by the voice assistant can be used for positive purposes only. In addition, we investigated and identified expected notification modalities in different contexts: 1) Users mostly prefer visual and application-based privacy notifications when interacting with their own voice assistant. 2) Users prefer audio and application-based privacy notification while interacting with classmate's voice assistant. Lastly, based on the responses from the user study, we implemented and evaluated the prototype for an audio-based privacy aware notification system to validate our conclusions of the preferred notification modalities. Evaluation of the prototype shows that the implemented audio and beep notifications are helpful, noticeable, and necessary to draw attention of the user.

While the results presented in this paper aids in advancing the discussion on privacy awareness in VUI devices, we propose some improvements which can further elevate the findings. The user study in this work was conducted in a lab settings with limited number of participants. Hence, we need to scale up the user research with more participants, to obtain statistically significant results. Furthermore, we could add another dimension to the qualitative analysis by critically relating the behavior and non-verbal expressions with the spoken words during interviews. To categorize and analyze the emotional experiences of a user, we have used Plutchiks wheel of emotions, which defines four primary emotions: joy, trust, fear, and surprise [20]. In order to obtain more granularity in understanding the emotional experiences and expectations, an additional analysis method could prove beneficial. The potential improvements of the current work are left for future investigations.

IV. Conclusion

Through the results presented in this work, we hope to have advanced the discussion on *privacy awareness of VUI devices*. By conducting user studies, and qualitative and quantitative analysis of the data, we present the findings on emotional experiences and privacy expectations of a user in a human-to-VUI interaction. The analysis shows that people are most concerned about privacy with respect to location tracking and listening to private conversations. We also discover the privacy expectations of a user from a voice assistant; they are: 1) a voice assistant should take consent for the usage of private data, 2) a voice assistant should have a feature to forget private data from users, 3) a voice assistant should support a feature to turn off the voice assistant, and 4) a voice assistant should only use the collected private data for positive purposes. Furthermore, we identify the user preferences for privacy notification modalities, with respect to different contexts. The results presented in this paper can be employed to design privacy-aware VUI devices, whereby user experiences and expectations can be modelled to allow the VUI devices to naturally adapt to the varying privacy requirements of the users without explicit user intervention.

References

[1] C. Jensen and C. Potts, "Privacy policies as decision-making tools: an evaluation of online privacy notices," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 2004, pp. 471–478.

[2] A. M. McDonald and L. F. Cranor, "The cost of reading privacy policies," *Isjlp*, vol. 4, p. 543, 2008.

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

[4] F. Schaub, R. Balebako, A. L. Durity, and L. F. Cranor, "A design space for effective privacy notices," in *Eleventh Symposium On Usable Privacy and Security ({SOUPS} 2015)*, 2015, pp. 1–17.

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

[6] J. Su, "Why Amazon Alexa is always listening to your conversations," https://www.forbes.com/sites/jeanbaptiste/2019/05/16/why-amazon-alexa-is-always-listening-to-your-conversations-analysis, 2019, accessed Nov 19, 2019.

[7] S. B. Wicker and D. E. Schrader, “Privacy-aware design principles for information networks,” *Proceedings of the IEEE*, vol. 99, no. 2, pp. 330–350, 2010.

[8] D. Coldewey, “This familys Echo sent a private conversation to a random contact,” https://techcrunch.com/2018/05/24/family-claims-their-echo-sent-a-private-conversation-to-a-random-contact, 2019, accessed Nov 19, 2019.

[9] A. Easwara Moorthy and K.-P. L. Vu, “Privacy concerns for use of voice activated personal assistant in the public space,” *International Journal of Human-Computer Interaction*, vol. 31, no. 4, pp. 307–335, 2015.

[10] F. Yeasmin, “Privacy analysis of voice user interfaces,” https://www.researchgate.net/publication/342110817_Privacy_Analysis_of_Voice_User_Interfaces, 2020.

[11] L. Stark, “The emotional context of information privacy,” *The Information Society*, vol. 32, no. 1, pp. 14–27, 2016.

[12] Mozilla, “Exploring the emotions of security, privacy and identity,” https://blog.mozilla.org/netpolicy/2013/05/21/exploring-the-emotions-of-security-privacy-and-identity/, 2019, accessed Nov 20, 2019.

[13] J. R. Reidenberg, N. C. Russell, A. J. Callen, S. Qasir, and T. B. Norton, “Privacy harms and the effectiveness of the notice and choice framework,” *ISJLP*, vol. 11, p. 485, 2015.

[14] R. L. Finn, D. Wright, and M. Friedewald, “Seven types of privacy,” in *European data protection: coming of age*. Springer, 2013, pp. 3–32.

[15] D. A. Norman, “Human-centered design considered harmful,” *interactions*, vol. 12, no. 4, pp. 14–19, 2005.

[16] J. W. Creswell, “Mixed-method research: Introduction and application,” in *Handbook of educational policy*. Elsevier, 1999, pp. 455–472.

[17] foodrisc, “Mixed methods research,” http://resourcecentre.foodrisc.org/mixed-methods-research_185.html, 2019, accessed Nov 20, 2019.

[18] J. Caulfield, “Thematic analysis,” https://www.scribbr.com/methodology/thematic-analysis/, 2019, accessed Nov 19, 2019.

[19] J. Rust and S. Golombok, *Modern psychometrics: The science of psychological assessment*. Routledge, 2014.

[20] R. Plutchik, “The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice,” *American scientist*, vol. 89, no. 4, pp. 344–350, 2001.