L. Hernández Acosta. and D. Reinhardt. Enhanced Privacy for Voice-Controlled Digital Assistants. Proceedings of the 20th International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom PhD Forum), pp. 130-131 2022.

Enhanced Privacy for Voice-Controlled Digital Assistants *

Luca Hernández Acosta
supervised by Delphine Reinhardt
Institute of Computer Science
Georg-August-University of Göttingen
Göttingen, Germany
hernandez@cs.uni-goettingen.de

***Abstract*—Voice-Controlled Digital Assistants (VCDAs) are increasingly used in both private and professional environments. While the convenience of use and the broad area of applications increase user adoption, the use of VCDAs involves a trade-off between user experience and privacy. In our research, we focus on developing solutions to better protect users' privacy by helping them to control the data shared with VCDAs.**

***Index Terms*—privacy, voice-controlled digital assistants, transparency, data control**

I. Introduction

The continuous development and improvement of voice processing technologies have increased the usability of Voice-controlled Digital Assistants (VCDAs) as well as their adoption in private and professional environments [1]–[3]. It is expected that the number of integrated VCDAs will reach up to eight billion devices in 2024 [4] with Amazon and Google currently leading the VCDA market. They are commonly used for, e.g., listening to music, asking questions, and checking the weather [5].

Data transferred to VCDAs, including audio recordings and information associated with users' accounts (name, address, payment information), are however stored by default on servers maintained by manufacturers or third-party developers. The audio recordings contain the users' voice commands, thus revealing their potential interests. Other characteristics about the users, such as their age, gender, or health status can further be inferred by analyzing their voice [6]. As a result, both manufacturers and third-party developers have access to fine-grained user profiles.

In order to protect users' privacy, manufacturers are proposing limited privacy measures. Firstly, users can mute their device to prevent recordings. Secondly, they can check their stored audio recordings and delete them manually or set automatic intervals for deletion if they have set up the device. Thirdly, they can create voice profiles, so that the device can distinguish them from other users. Last but not least, users can define and check the access permissions provided to third-party apps.

However, users are often not aware that these measures exist, or might find it too complicated to adjust them according to their preferences [7], [8]. Moreover, the available privacy measures offer only limited functions and do not cover all aspects included in the concept of privacy.

To bridge this gap, different solutions have been proposed in research. To prevent inference of user characteristics from voice analysis, a voice obfuscation method has been proposed in [9] to hide voice characteristics. Voice commands can further be cleaned from sensitive words, such as names or locations, as proposed in [9]. Risk traffic shaping [10] can be applied to obfuscate the original traffic that can reveal users presence at home, for example. Additional offline processing solutions methods, such as Snips [11] and Rhasspy [12] have been proposed to prevent the disclosure of the collected data to manufacturers and third parties. Finally, a solution based on continuous authentication [13] prevents unauthorized access to, e.g., dates or reminders by other people than the users themselves when interacting with VCDAs.

While these proposed solutions contribute in improving the users' privacy protection, additional efforts are necessary. For example, users have still only limited control over the collected, processed, and stored data. Moreover, there is a lack of transparency about these processes. As a result, our goal is to develop solutions that allow users to choose what data about them is stored and support them in making a decision by providing them usable transparency.

In the remainder of this paper, we hence detail our planned contributions in Sec. II and make concluding remarks in Sec. III.

II. Planned Contributions

Our goal is to develop a solution that allows users to better control which information about them is made accessible to manufacturers and third-party developers. As a result, we focus on threats to privacy coming from honest-but-curious entities that are part of the ecosystem and hence not on malicious attackers. As shown in Fig. 1, end users should be able to select which information is filtered prior to online processing and receive assistance in determining which records contain sensitive information and are considered for deletion. To reach this goal, usable transparency about the data processing should be provided to allow users to make informed decisions.

*The final publication is available at IEEE Xplore via http://dx.doi.org/10.1109/PerComWorkshops53856.2022.9767384

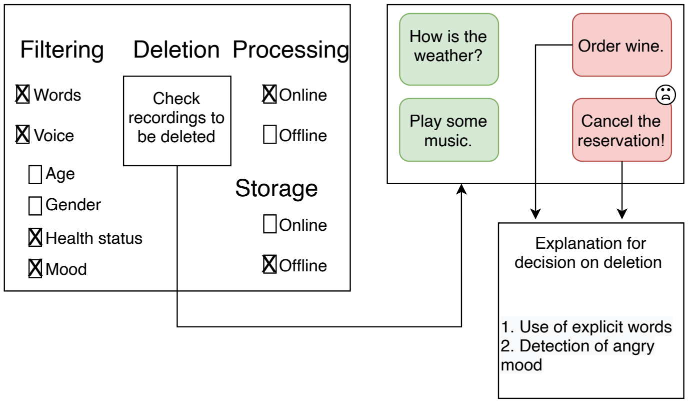

Fig. 1. Control Panel

Ideally, storing and processing data locally on the VCDA would prevent honest-but-curious entities from having access to user information. However, this is currently not feasible due to the limited devices' resources, functionalities, and increased user efforts that would negatively impact user adoption. Therefore, our work aims to contribute to the following four dimensions.

(1) **Data minimization**: Following the General Data Protection Regulation (GDPR) Art. 5 (1)(c) principle of *data minimization*, we want to reduce the collection of data irrelevant for providing voice assistance. Account information such as name, address or payment information should therefore not be stored or linked to users by the considered entities.

(2) **User-defined filtering**: Voice recordings are, however, desired for some users to view at a later time [7]. Still, those recordings might contain sensitive information that users want to filter out or delete afterwards. Therefore, it is required to explore different criteria for deleting sensitive voice recordings such as sensitive words, voices of children/ guests [7], or voice characteristics such as gender, health status, or mood. While some of these criteria have already been highlighted in related user surveys [7], [8], none focuses on sensitive voice content and information that can be inferred from the voice characteristics. Based on the obtained results, we will design a solution according to the criteria identified as relevant by the users to be filtered and not transferred to the concerned entities.

(3) **Sensitivity-based online/offline processing:** Furthermore, users should be able to choose whether a command should be processed offline or online, depending on their individual privacy preferences and how sensitive they perceive it. For example, if other smart home devices are controlled via voice commands, but the user does not want the manufacturers to know what kind of devices are in the user's home, they might prefer to process these commands offline. Even though offline processing is still a tedious task, this option could be available for user-defined sensitive commands.

(4) **Assisted selection of recordings to be deleted**: In order to not only automatically delete audio recordings, users should receive personalized assistance to understand why certain recordings could be deleted. When reviewing the record history, we will design a solution based on text and voice analysis with which sensitive records will be identified. The results will be presented to the users using a color code based on their sensitivity and accompanied by an explanation about its rating to support explainability.

III. Conclusion

Our goal is to propose a usable approach that provides more control to users about the processing and storage of their data collected by VCDAs. To achieve this goal, we will conduct user surveys to gather users' requirements and attitudes as well as evaluate the usability of our proposed solutions. Moreover, we will design and implement solutions that allow (1) to decouple the voice recordings from other account data, (2) users to define which features or commands are sensitive for them and should not be stored, (3) users to define which data should not be processed online based on their preferences, and (4) users to easily identify the recordings that could be deleted.

References

[1] M. B. Hoy, "Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants," *Medical Reference Services Quarterly*, vol. 37, no. 1, 2018.

[2] M. Porcheron, J. E. Fischer, S. Reeves, and S. Sharples, "Voice Interfaces in Everyday Life," in *Proc. of the 18th Conf. on Human Factors in Computing Systems (CHI)*, 2018.

[3] Z. Y. Chan and P. Shum, "Smart Office: A Voice-Controlled Workplace for Everyone," in *Proc. of the 2nd Int. Symposium on Computer Science and Intelligent Control (ISCSIC)*, 2018.

[4] Statista. (2021) Number of Digital Voice Assistants in Use Worldwide From 2019 to 2024. [Online]. https://www.statista.com/statistics/973815/worldwide-digital-voice-assistant-in-use/, accessed in 2021-12-22.

[5] B. Kinsella. (2021) Smart Speaker Use Cases. [Online]. https://voicebot.ai/2020/05/03/streaming-music-questions-weather-timers-and-alarms-remain-smart-speaker-killer-apps-third-party-voice-app-usage-not-growing/, accessed in 2021-12-22.

[6] J. L. Kröger, O. H.-M. Lutz, and P. Raschke, "Privacy Implications of Voice and Speech Analysis–Information Disclosure by Inference," in *Proc. of the 14th IFIP Int. Summer School on Privacy and Identity Management (IFIP SC)*, 2019.

[7] N. Malkin, J. Deatrick, A. Tong, P. Wijesekera, S. Egelman, and D. Wagner, "Privacy Attitudes of Smart Speaker Users," *Proc. on Privacy Enhancing Technologies (PoPETs)*, vol. 2019, no. 4, 2019.

[8] J. Lau, B. Zimmerman, and F. Schaub, "Alexa, Are You Listening? Privacy Perceptions, Concerns and Privacy-Seeking Behaviors with Smart Speakers," *Proc. of the ACM on Hum.-Comp. Interact.*, vol. 2, no. CSCW, 2018.

[9] J. Qian, H. Du, J. Hou, L. Chen, T. Jung, X.-Y. Li, Y. Wang, and Y. Deng, "Voicemask: Anonymize and Sanitize Voice Input on Mobile Devices," *arXiv preprint arXiv:1711.11460*, 2017.

[10] J. Liu, C. Zhang, and Y. Fang, "Epic: A Differential Privacy Framework to Defend Smart Homes Against Internet Traffic Analysis," *IEEE Internet of Things Journal*, vol. 5, no. 2, 2018.

[11] A. Coucke, A. Saade, A. Ball, T. Bluche, A. Caulier, D. Leroy, C. Doumouro, T. Gisselbrecht, F. Caltagirone, T. Lavril *et al.*, "Snips Voice Platform: An Embedded Spoken Language Understanding System for Private-by-Design Voice Interfaces," *arXiv preprint arXiv:1805.10190*, 2018.

[12] Rhasspy. (2019) Rhasspy Voice Assistant. [Online]. https://rhasspy.readthedocs.io/en/latest/, accessed in 2021-12-22.

[13] H. Feng, K. Fawaz, and K. G. Shin, "Continuous Authentication for Voice Assistants," in *Proc. of the 23rd Annual Int. Conf. on Mobile Computing and Networking (MobiCom)*, 2017.