Speech Triggered Mobility Support and Privacy

Michael Zipperle, Marius Becherer, Achim Karduck

Speech Triggered Mobility Support And Privacy

Michael Zipperle[1 2], Marius Becherer[1 2], and Achim Karduck[1]

[1] Furtwangen University, Furtwangen, Germany
[2] University of New South Wales, Canberra, Australia
michael@zipperle.de, marius.baech@gmail.com, karduck@hs-furtwangen.de

Abstract. Current voice assistants are offered by large IT companies such as Google, Amazon, Microsoft, Apple or Baidu. The voice assistants include numerous functionalities, which are usually executed centrally in the cloud by the providers. Nevertheless, the providers offer imprecise information on what happens to the input data of the users. Users cannot be sure whether their privacy and data are protected. The central research question is what is currently happening with the voice-based interaction between users and services, and what concepts for configurable data protection by users are conceivable in the future. In this article, we present the survey results obtained by speech assistant users. The results show, in particular, the willingness to pay for individually configurable privacy. The concept for a voice assistant with privacy-awareness is proposed and prototypically implemented.

Keywords: Data Security, Privacy, Voice Assistant, Cloud Computing, Natural Language Processing, Mobility Support

1 Introduction

Voice control is an interaction possibility where services can be controlled by human speech. Realizing voice control requires an ecosystem - the voice assistant - out of several components to deal with tasks such as converting in both ways text to speech, speech to text as well as extracting metadata from text input. Hereby, several processes can be initiated such as playing the next song, setting the alarm clock or starting an ordering process. Experts expect a growing market for voice assistants: The trade journal "PR Newswire" assumes that purchases via speech will increase twenty-fold in the next four years [5]. The magazine "Campaign" estimates that in the future the search in browsers using the keyboard will be replaced by the search via voice [15]. The voice assistants preprocess the voice in order to understand the meaning, and therefore, control several services regarding the voice input. For this reason, the voice assistant can be considered as a component between user and application that enables the interaction by the human voice.

The applications of a voice assistant run on a platform in the cloud. Speech processing on the platform is demanding because a user's voice input runs through highly complex sub-processes of speech processing. It takes time until a suitable user response is generated. Currently, these platforms are offered by major cloud providers who have the computational resources and know-how of each sub-process. Voice assistants are offered by Amazon, Google, Microsoft or Baidu, with many functionalities and excellent performances. However, there are privacy concerns as the privacy policies of cloud providers do not clearly state what happens to users' data in the cloud. Mobile devices such as smartphones and speakers send a user's voice input to the appropriate cloud provider for analysis. The data processing is non-transparent and it opens up the opportunity to establish a side-channel between the user and its applications. In the process, data can be collected and possibly misused. Nowadays, sensitive data are shared among users and enterprises, for instance, the banking system, health care system and messaging. The use of data is specified in terms of use, but these give a limited indication of the possible usage scenarios, such as user profiling.

In this work, we conduct a survey with 110 participants to understand today's customer needs of their perspective on privacy in the domain of voice assistants. Regarding the survey, we develop a privacy-awareness concept that includes user-controlled privacy with a high level of functionality as well as performance. Besides, the concept is implemented into the privacy-awareness ecosystem consisting of three separate components: the mobile app, privacy provider, and voice assistant.

The results explained in section 3 represent the motivation for the development of a privacy-awareness concept for a voice assistant. In section 4.1, the privacy-awareness concept is presented to transfer full control back to the user. A full-stack architecture is shown in section 4.2. Then, in section 5, a prototype is presented that implements this architecture. The presented prototype allows maximum flexibility in data control and provides the user with configurable pri-

vacy. A summary of the concept, the technologies, and the developed prototype conclude this article.

2 Related Work

Voice assistants such as Amazon Alexa, Google Assistant, Apple Siri, Microsoft Cortana, and Baidu DuerOS currently dominate the market. However, they all are non-transparent in their voice assistant ecosystem and providing a minimal declaration of data usage [2,7,4,12,9]. In detail, the services in the cloud infrastructure are a hidden secret, and therefore, less privacy-awareness are provided for the users.

The conducted study on privacy risks of voice assistant apps presents diverse risks in the communication, the user identification scheme, as well as privacy-related information [13]. During their investigations, they expose that Google Assistant transfer several additional information, such as user information and user device information. Furthermore, in 50.51 % of the voice assistant apps, linking customers to an identification scheme has been feasible. More significant, personal information such as birth date, name, e-mail address, blood type, gender phone number are tried to obtain from the apps. This demonstrates the high usage of data capturing data in the case of Google Assistant. Likely, the same is happening in other commercial voice assistants.

Even though the previous study exposes exploited data collection, many customers are not conscious of these hidden processes. An investigation on smart home user's indicates lacking conscious of their privacy perception of the devices, and consequently, they prioritize convenience over privacy [17]. Another study reveals the incompleteness of data activities to user perception that leads to design implications in the domain of smart home voice assistants [1]. In the security and privacy domain of voice assistant, user perception with their concerns is conducted and presents less customer disposition in discussing the costumer's voice assistants [6]. In contrast to the studies that cover users' perception and knowledge, the authors of [11] conducted a study with users and non-users of voice assistants and compared their reasons that support and criticize voice assistants. Hereby, customers did not perceive different mechanisms such as audio logs as privacy control, and additionally, other privacy controls often are not used.

Despite the forgone studies in this domain, most of them focusing on user perception and conclude lacking privacy awareness. The one weakness of those studies is the limited user feedback regarding the conducted interviews, but the main issue relies on lacking user expectations. Hereby, the studies do not understand the user needs and there willingness to overcome those privacy concerns.

Different approaches to deal mitigate privacy risk such as an access control mechanism that is used in the smart home domain by with information are delivered regarding the task context [8]. In a more abstract representation, fundamental design strategies are proposed that cover data economy, careful allocation, user ability to control, and usability of security mechanisms [16]. More

specific details in terms of privacy are conceptualized in the privacy framework with a focus on privacy key risks such as system development and design and cross-organizational collaboration [14].

Overall, the previous research supports the assumption that data is collected even though not every time permission is required, such as for device ID and user ID. Consequently, users cannot control data access and monitor the data processing in clouds of global players. Although the majority of customers have a significant knowledge deficit in the privacy perception of voice assistants, the studies do not address the personal perception of specific data. Without knowing the customer's need for privacy, it is challenging for developers to understand what data violates the individuals' privacy. Furthermore, several privacy frameworks have been proposed as well as concrete implementation in access control mechanisms for task-oriented service computing. Whereas privacy frameworks are imprecisely but user-centric, concrete access control mechanisms are precisely but service-centric. Ultimately, the current research is lacking a privacy-awareness concept that considers configurable users' privacy with a large extent in functionality.

3 Motivation

In the beginning, an initial survey was carried out to determine whether more data protection is desired for voice assistants in Germany. Over one hundred participants took part in the online survey. 54.5 % of the participants were male, 45.5% female, and 10% were under 18, 69% between 19 and 25, 7% between 26 and 35, and 14% over 35 years old. The participants were asked the following questions:

1. How often do you use a voice assistant?
2. Do you know what happens to your data?
3. Which particularly privacy-related information is important to you?
4. Would you pay money for high privacy?
5. How much money would you pay once for high data protection of an application?

The first question revealed that 44.5% use a voice assistant once a month or more often, whereby 90% of the participants do not know what is happening with their data. The results show that a high level of privacy is desired especially for their banking credentials, individual behavior and attitudes, communication among individuals, contacts, interests, and location.

One in four would pay for better privacy, and 56% of the participants are unsure if they would spend money on it. The group under the age of 18 years has the least willing to pay. The intersection of participants who ticked "Yes" or "Maybe" increase with age. The amount that participants would spend on better privacy varies widely. Approximately 15% of the respondents were not willing to pay for it, while the majority are. 50% of the respondents would pay up to 5€, 15% 6 - 10€ and 20% over 10€.

Therefore, the following conclusions can be drawn from the survey results: First, voice assistants are used to varying degrees and users do not know what happens to their data. Second, privacy for users is essential, even though the exposure level of data regards to individuals' perception and third, users would pay for the protection of their data.

4 Methodology

This section proposes a concept for high privacy-awareness, and the architecture to implement it.

4.1 Concept

Based on the survey, a concept for a voice assistant was developed. The costs for the required resources were neglected. An important requirement of the concept is the privacy-awareness considering the design principles for the multilateral privacy according to Kai Rannenberg focus on the following four points[16] and the privacy framework [14]: Data minimization, control possibilities for the user, possibilities of choice and room for negotiation, and decentralisation and distribution.

Within this concept, the focus is on the first three points. Often applications collect data from a user, which the user did not agree with, for instance, sharing device id or user information [13]. Therefore, the concept aims to ensure that an application only collects data from users who need it. Furthermore, the provided data for application should be, both physical and operational, fully-controlled by the users to enable privacy-awareness data access. By doing so, users have the opportunity to decide if they want to share no data, partially data, or the full requested data. Since some application depends on the requested data to work properly, data can be manipulated in the middle by the user to ensure privacy.

User-controlled privacy allows a user to determine what data he or she releases for specific applications. However, applications require additional data from a user to provide sufficient usability. An example is a voice assistant's question about weather forecasts. If the voice assistant knows the user's location, it can provide the weather forecast for the user's position. Otherwise, the voice assistant would first have to ask the user for which location he or she wants a weather forecast. If a user does not want to release his data for an application, he can define a fictitious context. This allows the user to use this application, but at the expense of lower user-friendliness.

4.2 Architecture

The concept for high privacy-awareness can be implemented using the architecture and technologies shown in figure 1. Three components are required for the implementation: First, the Mobile App serves as an interface to the user. The user can make voice entries and configure his profile and privacy via the

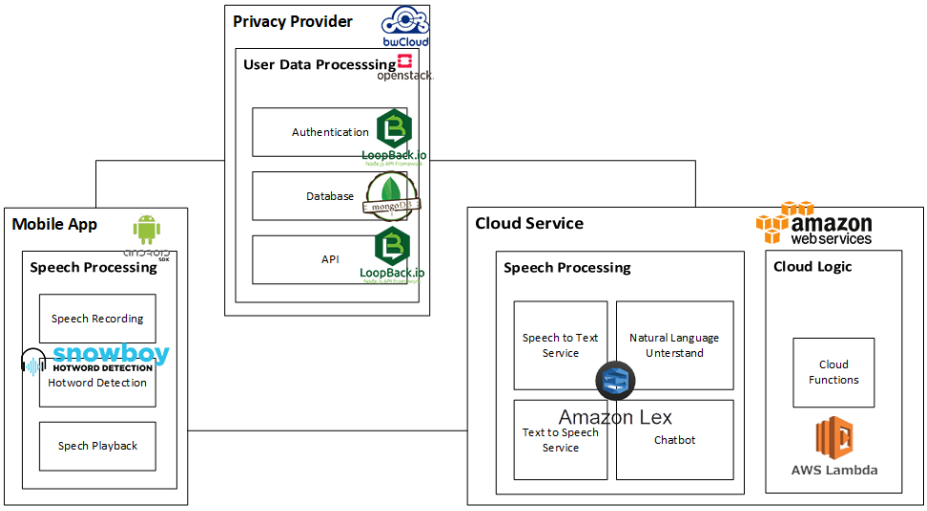


Fig. 1: Architecture Overview

app. Second, cloud services handle voice processing and generates output to the user. The privacy provider is accessed to generate a response. Third, the Privacy Provider is the core component, which guarantees the user more privacy. The exact structure and choice of technology are described below.

5 Prototype

In the following, the application example is described, and then the implementation of the individual components of the architecture is explained in more detail.

5.1 Application Example

The application example is intended to show how a voice assistant can be used to promote speech-based mobility support and at the same time, ensure fine-grained data protection for the user. The application example covers the following functionalities: First, the search for doctors can be carried out in a specific location. Secondly, a doctor's appointment can be negotiated, whereby a doctor's calendar is compared with the user's calendar. Finally, optional mobility support such as scala mobile, assistance to the user when leaving the house or a pick-up service, that takes the user to the doctor and back home, can be requested.

A user can define the data required for these functionalities in detail in a mobile app. For example, a user can have his location determined automatically via GPS to expose his current location, but he can also edit the location manually to preserve it location. The user can share his calendar, but no appointment details are necessary for the application example. It is sufficient to know whether

the user is available at a particular time or not. For this reason, a user can hide appointment details like title and description of an appointment. More details about privacy will be shown in the course of this section.

5.2 Privacy Provider

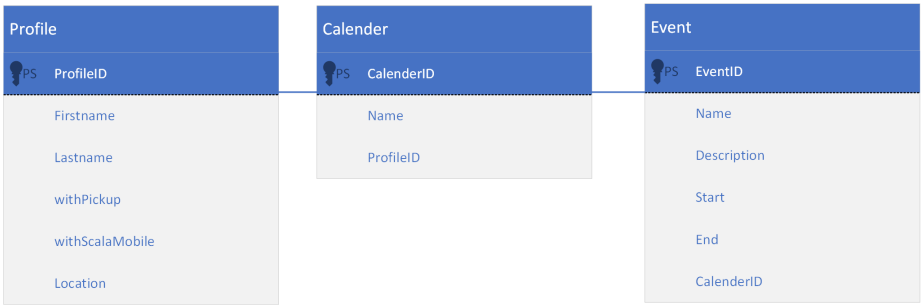


Fig. 2: Data Model for the Application Example

Information on the user context is to be stored in the privacy provider provided. The data model is explicitly described for the application example. This data model describes the user context only to a minimal extent. However, the data model can be extended. In this data model, which is shown in figure 2, the user can select different profiles. This way the correct name can be given or the identity can be hidden behind another name. In the profile, information about mobility is stored, e.g. whether a driving service and a Scala Mobile is required. Calendars are linked to the profile data. The data is read directly from the smartphone, but other online calendars can also be added. Finally, a calendar contains several events. For the location, there is the location field. The location can be determined via the GPS sensor, or a location can be entered manually.

5.3 Mobile App

The mobile app serves on the one hand as an interactive interface between the user and the voice assistant and on the other hand, as a privacy configurator. The app offers the following four views:

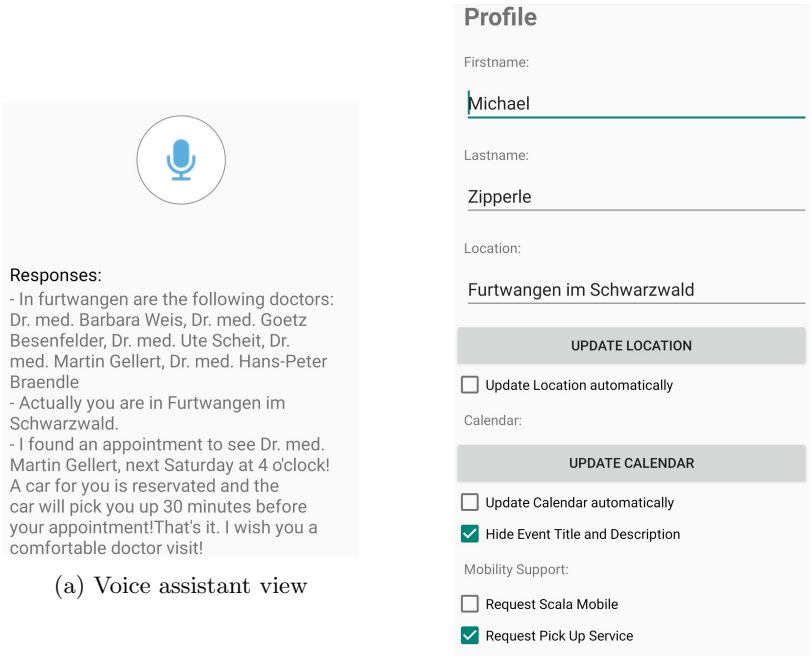
- Registration: A new user can register with the voice assistant with his e-mail and a password. The user can choose between the profile types private, doctor or mobility support. Depending on the profile type, different privacy settings are preconfigured. By default, a private profile is only accessible to the owner, whereas the profile types Physician and Mobility Support are publicly accessible.
- Login: A registered user can log in to the voice assistant with his e-mail and password.

- Voice Assistant: The voice assistant is the main view of the app, which appears as soon as a user is logged in. At the same time, the Hotword Detection is started, which was realized with Snowboy from Kitt.ai [10]. The Hotword Detection listens locally on the Smartphone until the signal word "Butler" is recognized; no data is stored and passed on to third parties. After detection of the signal word, voice processing is moved to the cloud for better performance. As soon as an interaction between user and voice assistant is finished, the hotword detection is reactivated. Also, all the output of the voice wizard on the view. The audio recording and playback were implemented with the Android SDK.
- Settings: This view allows the user to configure his profile and privacy settings. The first and last names can be set for the profile. The location can be determined either by text input or by GPS. When determining the location via GPS, the user can decide whether the location is retrieved once or whether the app can automatically update. The user can make his calendar available to the app. Besides, the automatic update of the calendar or the hiding of details such as the title and description of an appointment can be activated. This ensures that only the data relevant to the application example is accessed. Other personal data, such as the name of the calendar or the owner, cannot be determined by the app. Next, the user can define mobility supports that are requested by default when an appointment with a doctor is made. On the one hand, Scala Mobile, which helps the user to leave the house, or on the other hand, a pick-up service, which takes the user to the doctor, can be requested. The profile and privacy settings are synchronized with the privacy provider. This app view enables the data economy and user-controlled privacy presented in the concept.

5.4 Cloud Service

As soon as the hotword detection has detected the signal word "Butler", the further processing of a user's voice input is performed in the cloud. By outsourcing resource-intensive voice processing to the cloud, users can be assured of high performance and ease of use. For the realization the following Amazon Web Services (AWS) were used:

- Amazon Lex: Amazon Lex serves as a conversation interface for speech and text. Based on the text intent, an output can be generated, which is then converted from text to speech[3]. An Amazon Lambda function is triggered to generate a response to a text intention.
- Amazon Lambda: Based on Amazon Lambda, an application was developed that generates a dynamic response based on the text intention of the user. The application can access the privacy provider to retrieve the necessary information to generate the response by using the user's access token. Thus, the applications could realize the negotiation of a free doctor's appointment between doctors and user.



(a) Voice assistant view

(b) Preference view

Fig. 3: Mobile App

5.5 Data Release

To use data from the privacy provider, the user must first be authenticated. To do this, the registration data is entered into the mobile app. When the login button is clicked, the user is authenticated by the privacy provider. During this process, the user is granted read, write and execute rights, so that data can be changed and various functions can be performed. Besides, a time-limited key is generated with which the user can grant applications access to his data. Which data can be read by an application can be released by the user in the mobile app.

6 Conclusion

Some conclusions could be drawn from the developed prototype, which offers high user security and control possibilities. It was possible to take the concept into account during implementation and aspects of multilateral security, and user-controlled privacy is also found in the prototype. The prototype consists of the voice processing environment (cloud services), the mobile app and the privacy provider.

The functionality requirements were met using Amazon's cloud services. Besides, the fulfilled General Data Protection Regulation (GDPR) guidelines of the voice services can guarantee data protection for the users. By using the

speech-based cloud services, developers do not have to deal with speech processing in detail but can develop an application on an abstract level. Before a user can use the voice assistant, he or she must authenticate. This protects access from unauthorized persons. In the app, users can create different profiles with different data, the use of pseudo profiles is possible. Concerning the concept of multi-layered security, this is important to create choice and room for negotiation for the user. The user data is stored in a privacy provider. The concept of multi-level security also applies to the privacy provider. Decentralization and distribution are of great importance here. The choice of technology and provider takes data protection into account at all levels. However, there is still potential for optimization at the privacy provider, resource access should be made more configurable through authorization, duration, and filtering.

Depending on the user, the requirements for a voice assistant vary. Different applications should be able to be activated or deactivated based on the needs of a user. This functionality could extend the prototype in the future. The applications offered must inform the users about user data and thus create transparency. A standard for data storage must be created to create an ecosystem in which every application can access the data. Otherwise, the applications will have to manage the user data themselves, and the concept of separating data and application would become obsolete.

In this article, the potential of voice assistants is referred to at various points. By the pleasant handling of the system intelligence, it offers an added value in everyday life. However, different industries must open up and offer interfaces so that bookings and reservations are not only possible by e-mail or telephone. If the infrastructure of companies is created, voice assistants will become even more attractive for users.

References

1. Abdi, N., London, C., Ramokapane, K.M., Such, J.M., London, C., Clara, S.: More than Smart Speakers : Security and Privacy Perceptions of Smart Home Personal Assistants This paper is included in the Proceedings of the (2019)
2. Amazon: Alexa internet privacy notice (2018), https://www.alexa.com/help/privacy
3. Amazon: Amazon lex (2018), https://aws.amazon.com/de/lex/
4. Apple: Privacy policy (2018), https://www.apple.com/legal/privacy/en-ww/
5. Consultants, O.S.: Voice shopping set to jump to $40 billion by 2022, rising from $2 billion today (2018), https://www.prnewswire.com/news-releases/voice-shopping-set-to-jump-to-40-billion-by-2022-rising-from-2-billion-today-300605596.html
6. Fruchter, N., Liccardi, I.: Consumer Attitudes Towards Privacy and Security in Home Assistants. In: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18. pp. 1–6. ACM Press, New York, New York, USA (2018). https://doi.org/10.1145/3170427.3188448, http://dl.acm.org/citation.cfm?doid=3170427.3188448
7. Google: Privacy policies (2018), https://policies.google.com/privacy?hl=en-US

8. He, W., Golla, M., Padhi, R., Ofek, J., Dürmuth, M., Fernandes, E., Ur, B.: Rethinking access control and authentication for the home internet of things (iot). In: 27th {USENIX} Security Symposium ({USENIX} Security 18). pp. 255–272 (2018)
9. hemple, J.: How baidu will win china's ai race - and, maybe, the world's (2017), https://www.wired.com/story/how-baidu-will-win-chinas-ai-raceand-maybe-the-worlds/
10. Kitt.ai: Snowboy hotword detection (2018), https://snowboy.kitt.ai/
11. Lau, J., Zimmerman, B., Schaub, F.: Alexa, are you listening? Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. Proceedings of the ACM on Human-Computer Interaction **2**(CSCW) (2018). https://doi.org/10.1145/3274371
12. Mircosoft: Cortana and privacy (2018), https://support.microsoft.com/en-us/help/4468233/cortana-and-privacy-microsoft-privacy
13. Natatsuka, A., Akiyama, M., Iijima, R., Sakai, T., Watanabe, T., Mori, T.: Poster: A first look at the privacy risks of voice assistant apps. Proceedings of the ACM Conference on Computer and Communications Security pp. 2633–2635 (2019). https://doi.org/10.1145/3319535.3363274
14. NIST: Nist Privacy Framework: a Tool for Improving Privacy Through Enterprise Risk Management (2020), https://www.nist.gov/system/files/documents/2020/01/16/NISTPrivacyFramework{_}V1.0.pdf
15. Olson, C.: Just say it: The furture of search is voice and personal digital assistant (2016), https://www.campaignlive.co.uk/article/just-say-it-future-search-voice-personal-digital-assistants/1392459
16. Rannenberg, K.: Multilateral security a concept and examples for balanced security. Proceedings New Security Paradigm Workshop pp. 151–162 (2000). https://doi.org/10.1145/366173.366208
17. Zheng, S., Apthorpe, N., Chetty, M., Feamster, N.: User perceptions of smart home iot privacy. Proceedings of the ACM on Human-Computer Interaction **2**(CSCW), 1–20 (2018)