Without Being Asked: Identifying Use-Cases for Explanations in Interaction with Voice Assistants

Dominik Pins
Fraunhofer Institute for Applied Information Technology
Sankt Augustin, Germany
Dominik.Pins@fit.fraunhofer.de

Fatemeh Alizadeh
Department of Information Systems and new media
University of Siegen, Germany
Fatemeh.alizadeh@uni-siegen.de

ABSTRACT

Voice Assistants (VAs) support the user in coping with daily life. By controlling the device via speech, the interaction is subject to similar challenges as in human interaction. While in human interaction the people involved can usually clarify misunderstandings and unexpected behaviors that arise, VAs often lack solution-oriented explanations, which confuses the user. Through interviews with users of VAs, we identified three use cases that illustrate user confusion and require further explanation. The most common is accidental activation, followed by wrong functioning and execution by incorrect devices (when more than one device is used in a household). In this study, we aim to draw attention to confusing situations and contribute to designing human-friendly explanations that help users understand the behavior of VA and have more satisfying interactions.

KEYWORDS

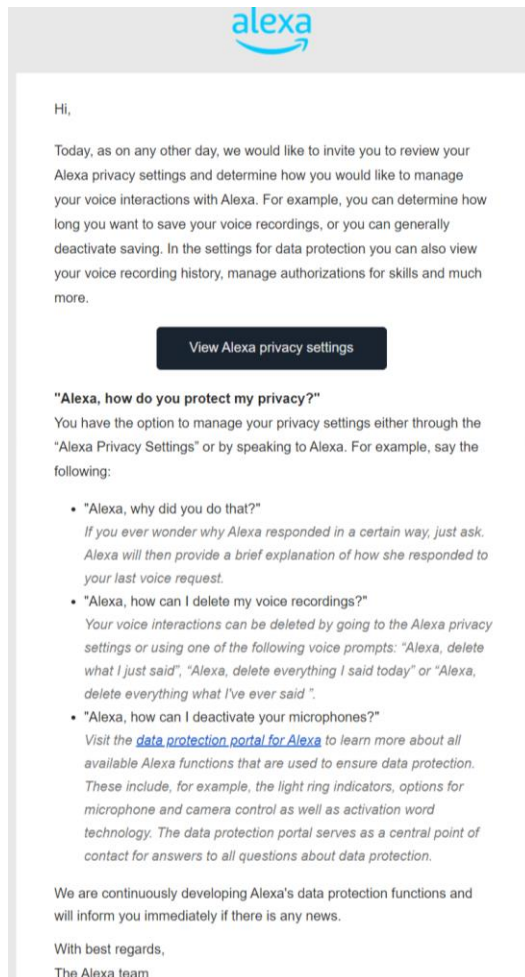Voice assistant, Explanations, Unexpected behavior, Misunderstanding

Figure 1: The email that Alexa users have received from Amazon after the new updates (the original email has been sent in German and was translated into English)

INTRODUCTION

Interaction between two objects, human or technical, is full of risks and misunderstandings. Voice assistants (VA) represent a particularly interesting case here, which are often attributed both technical and human characteristics due to its capabilities [20, 27]. On the one hand, the human-like interaction of the user with the VA enables various tasks via speech and offers its advantage mainly through hands-free use. On the other hand, this means that just like human-like communication (as shown by various communication models, e.g. Shannon and Weaver's model [32]) they are also prone to misunderstandings and misinterpretations.

Despite the increasing popularity of these devices, there is still little support from them to solve the communication breakdowns and the burden of ensuring a successful communication lies in human, who must adapt his communication patterns to the needs of the VAs [4].

The major challenge in the case of VAs is that, unlike human interactions where both parties can immediately ask questions to clarify possible misunderstandings, VAs often act in a less clearly solution-oriented manner and tend to end the interaction abruptly [26]. Previous studies have shown if communication malfunctions occur, users adopt various repair strategies to solve the problem. However, this process requires a lot of effort and is not always successful [4]. As a result, users of VAs are often confused or even frustrated when the VA does not behave in the expected way and they are unable to understand or explain why. To address this issue, Amazon recently added a feature through which users can request explanations from Alexa about the previous interaction by asking "Why did you do that?" [3]. As can be seen in Figure 1, this question is meant to give users more information when the device does something unintended, like setting a timer out of nowhere [24].

Explainability of systems not only has the advantage of helping users easily understand and analyze the actions of systems, but it can also be used to implement a social right to explanation [14]. However, researchers have argued that not every explanation is useful. According to Monlar, user-friendly explanations are those that explain anomalies, i.e., why a system does not behave as expected [22]. In other words, before allowing users to ask why, we need to find out the situations in which their expectations are violated by the system. In this sense, and on the way to user-friendly explanations when communicating with VAs, we aim to identify potential use cases for explanations of VAs and users' needs for more information about the system's whys.

RELATED WORK

Interaction with voice assistants such as Alexa (Amazon), Siri (Apple) or Google Assistant is a current topic in HCI research. This concerns firstly the use and "living" with a VA [5, 8, 26, 31, 33], secondly the design of the behavior of a VA and the interaction itself [6, 13, 15, 23, 28] and thirdly the privacy issues that come up with its use [1, 7, 9, 19]. To set a framework for this study, we first look at previous work on designing human-friendly interactions and why users need them and

then move the previous on difficulties in interaction with VAs. In this way, we highlight the research gap in mapping users' explanation needs to their challenges in communicating with VA.

Designing Explanations

Humans tend to acquire and provide explanations for various reasons. The one most often mentioned is "to predict similar events in the future"[16]. Explanations are also used to ask why a system failed in order to repair it for normal function [16]. But the algorithmic systems that play an ever-increasing role in our daily lives are becoming more opaque by the day [11]. Therefore, the lack of authorized explanations forces us to come up with our own theories (called folk theories) about how these systems work in order to interact with them and make decisions [10]. However, these folk theories are often incomplete and prone to misunderstanding [2]. To refine the user's mental model and eliminate these misunderstandings, explainable artificial intelligence (XAI) has been introduced as AI systems that users can easily understand and explore [14, 29]. However, researchers have argued that explanations are not always beneficial and too much transparency could also burden and confuse users [11]. Therefore, human-friendly explanations have been defined as those that explain anomalies. That is, when the system does not work as users expected it to [22].

In the past, voice assistants have raised some privacy concerns due to their constantly eavesdropping microphones [19, 25]. To compound the problem, new updates to the European General Data Protection Regulations [12] have also given users the right to explanations about the system's logic (see article 15 and 22 [12]). As a result, Amazon recently gave Alexa's users the ability to ask Alexa why it behaves in a certain way [24]. However, designing human-friendly explanations and providing the right level of transparency would mean that we need to figure out the right use cases first before allowing users to ask "why" so as not to burden users by complicating their interactions. So, what are the situations where users are looking for information and are confused by the lack of explanations in the context of interacting with voice assistants? In this study, we try to answer this question.

Challenges of Human VA Interaction

Interaction with voice assistants is often described by their users as pleasant, successful and useful [17, 26, 27]. Nevertheless, users also report accidental activations and unexpected behaviors of VAs that occur regularly [4, 26]. This was also shown in a study by Schönherr et al. [30], who investigated accidental triggers of VAs by similar wording of the activation words.

In the same vein, Lahouhal et al. [18] investigated how users behaved in situations where the VA did not understand the command correctly or behaved unexpectedly, and identified seven reasons why users of VAs forgave and accepted the errors (the so-called usage paradox). In their study, among other reasons, users' understanding of what did not work correctly led to acceptance of the errors in the interaction. In a previous study we could already point out that users stop using a function such as cooking with a VA when the complex interactions are error-prone and the VA

#	Age	Sex	VA system	In use since
P1	35	M	Amazon	Jul'17
P2	27	M	Google	Oct'16
P3	48	M	Amazon	Nov'16
P4	26	F	Amazon	Jul'19
P5	35	M	Amazon	Jul'17
P6	24	M	Amazon	May'19
P7	31	M	Amazon	May'18
P8	56	M	Google	Jul'20
P9	34	M	Amazon	Dec'17
P10	58	F	Amazon	Jul'18
P11	34	M	Amazon	Dec'19
P12	32	F	Amazon	Sep'20
P13	31	M	Amazon	Sep'20
P14	17	M	Google	Sep'20
P15	50	M	Amazon	Sep'20
P16	27	M	Google	Dec'19
P17	34	M	Amazon	Feb'19

Table 1: Participants' information of our study

cannot provide feedback or suggest a solution [26]. Therefore, insufficient feedback about what the VA does and how it does it, leads to limited user interaction with it [21]. In this work, we aim to find out whether users seek explanations in these situations and what other confusing situations exist within the interaction with VAs.

METHODOLOGY

Following the semi-structured interview method, we conducted 17 in-depth interviews with VA users (14 m, 3 f), with an average age of 35, and asked them about their use of VA's and their privacy concerns. Participants were all German-speaking and were recruited through authors social networks and the snowball system, most of them being part of a Living Lab Study. Participants either lived in mixed households with their family, partners, roommates or alone. For the study, we focused on the account holders of the devices as they are usually also the main users. It turned out that the account holders we mainly talked to were usually the male members of the household, which was reflected in the analysis, but they also correspond to the general user group of smart home devices. Participants had used either Amazon Alexa or Google Assistant for between one month and four years. Due to the COVID-19 situation, the interviews took place remotely via a video conferencing tool. They lasted approximately one hour and were recorded with the participants' consent. In the analysis phase, the inductive analysis method was used, where two different researchers manually coded the interview transcripts. The results were assigned to six different themes, which were discussed in an online session until agreement was reached. Only the findings that relate to the research question of this study are described below.

RESULTS AND DISCUSSION

By analyzing the results, we identified three potential use cases where explanations are needed by VA users. We also recognized two different states that occur in these situations. As a result of the confusion, participants began to self-explain why the situation occurred. Either they were able to find a logical reason for VA's behavior, or they were unable to find it and continued to be frustrated and skeptical of VA. We argue that we can prevent this further frustration and distrust of the system by helping users understand confusing situations.

Unintended activations

"Sometimes Alexa starts by herself and then we get quieter to hear what she is doing and wait for her to finish speaking." (P10)

All of our participants reported situations in which their VA suddenly began to speak, even if it was not directly activated. This usually happened during a conversation or when they were watching TV. While some situations were easily resolved because participants were talking about the voice assistant or Googling something and therefore accidentally said the activation word, there were also situations where TV was not turned on or participants were talking about something irrelevant. In these situations, users were confused and tried to find a logical reason for

what was happening. One suspected reason was that VA might understand the activation word by similar words or 'accidental triggers' [30].

One participant noted that sometimes the "Listen" signal lit up, but then VA noted that there were no commands. This seemed to have improved in his opinion, "She used to react very strongly to the TV. Maybe now they've agreed on what advertising is and she doesn't respond to it, or she recognizes the sound quality." (P2)

Most confusing were situations where the user could not find an explanation for the activation, as one participant recounted, "We were wondering why Alexa starts talking upstairs in the bedroom while we are downstairs in the living room watching TV. We couldn't understand her, but there was some babbling." (P1)

In both cases, understandable or not, participants were irritated and wondered why unintended activations were occurring. However, they usually did nothing to clarify or change the situation. They simply let the VA talk undisturbed and then continued with whatever conversation or activity they were engaged in. "Sometimes something will activate her, even if you weren't talking to her at all. Then she says, 'I'm sorry, I don't know that,' or she tells you something and you think to yourself, 'Aha, but that's the information I didn't ask for.'" (P9) Only one participant indicated that he was annoyed enough by repeated unintended activations to consider doing something about it: "At first I'm surprised and when it happens more often, I get annoyed. I'm always on the verge of turning the microphone off." (P8)

Wrong functioning

"Sometimes my tongue slips and then I try to have a dialog with the machine. Then it gets pretty humanized, like: I said so and so, Alexa." (P3)

The failure of an interaction is another use case we identified during the interviews that often led to confusion among the participants. In these situations, the VA did not properly understand the speaker. In two cases, participants assumed that VA would be better able to understand users if they were account holders (mostly male users): "When my girlfriend approached the VA, it didn't work. I'm not sure if it's because I trained it with my voice once or not. But it was like she said something and nothing happened and then I tried it and it worked." (P7)

As shown in the opening quote, malfunctions can also occur because of a misunderstanding. For example, the VA sets an incorrect timer or misunderstands a song: "Sometimes when I try to play a certain song or something, she doesn't understand it. The problem is when you ask something in English and she only knows German, she doesn't understand the foreign words." (P4) Another participant also suggested that his VA might have problems with understanding (foreign) words: "I tried to find out if Google can tell me when I have my next hairdresser [in German Friseur] and foot care appointments, it didn't understand the word "Friseur" at all, but brought me all my foot care appointments. Somehow, I must not have set the appointments with the correct word, because unlike "foot care", "Friseur" is not originally a German word, but French" (P8)

As with accidental activation, sometimes the participant could not explain why something happened. For example, Google Assistant Users could link different accounts to the assistant to get

more personalized information and services. One participant who used this feature was confused when a guest asked his VA "Who am I?" and the VA recognized the guest as the owner of the device. He tried to find a logical reason for this error, but since the two voices were not similar, he was irritated. One participant suspected Alexa of adding an item to his shopping card. Although he was not sure if it could have been someone else in the household who added the item, he researched the Internet to find out whether or not another VA user had such an experience.

Execution of commands by the wrong device

"Our dining room is between our two VAs, the one in the kitchen and the one in the living room, so it's sometimes a bit tricky. Sometimes you have to speak in the direction of the kitchen door or speak loudly in the direction of the living room for the right device understands you." (P1)

Three participants told us that they had problems addressing the correct device. Two of them could somehow explain this misunderstanding with the physical proximity, one could not understand how this could happen even though the two devices were far away from each other "[...] you have to shout to control the device that is further away." (P17) One participant suggested that this problem could be solved by changing the settings regarding the priority or sensitivity of the device to respond to the activation word but had not yet tried where and if this was possible.

Conclusion and future work

The results so far show that users have assumptions and try to find explanations for the unexpected interactions with VA and become frustrated and sceptical when they do not find a logical reason. However, they rarely or never took steps to resolve these misunderstandings and avoid them in the future. This forgiving and accepting can also help them feel a sense of agency and control in the face of the VA unexpected behavior.

Our next steps are to ask participants in the Living Lab study to test the ability to ask Alexa (and Google Assistant) why they did the last execution in different situations after using it. In this way, participants can evaluate VA 's response and give us further clues about information needs and explanation expectations. In addition, we plan to hold a workshop on unexpected interactions to find out how unexpected interactions are defined by participants, what explanations VA should provide to clarify them, and what actions need to be taken. Furthermore, we want to discuss what data or information about the interaction could help further to support sense-making in these situations.

ACKNOWLEDGMENTS

REFERENCES

[1]   Abdi, N. et al. 2019. More than Smart Speakers: Security and Privacy Perceptions of Smart Home Personal Assistants. (2019), 17.

[2]   Amazon.com Help: Ask Alexa to Tell You Why She Just Did Something: *https://www.amazon.com/gp/help/customer/display.html?nodeId=GJT6X5TZUW8AB9Y9*. Accessed: 2021-02-24.

[3]   Beneteau, E. et al. 2019. Communication Breakdowns Between Families and Alexa. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow Scotland Uk, May 2019), 1–13.

[4]   Bentley, F. et al. 2018. Understanding the Long-Term Use of Smart Speaker Assistants. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. 2, 3 (Sep. 2018), 1–24. DOI:https://doi.org/10.1145/3264901.

[5]   Beschnitt, M. 2017. NEXT GENERATION USER INTERACTIONS: CONVERSATIONAL UX. (2017), 43.

[6]   Cho, E. 2019. Hey Google, Can I Ask You Something in Private? *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19* (Glasgow, Scotland Uk, 2019), 1–9.

[7]   Cho, M. et al. 2019. Once a Kind Friend is Now a Thing: Understanding How Conversational Agents at Home are Forgotten. *Proceedings of the 2019 on Designing Interactive Systems Conference - DIS '19* (San Diego, CA, USA, 2019), 1557–1569.

[8]   Chung, H. et al. 2017. "Alexa, Can I Trust You?" *Computer*. 50, 9 (2017), 100–104. DOI:https://doi.org/10.1109/MC.2017.3571053.

[9]   Hagras, H. 2018. Toward Human-Understandable, Explainable AI. *Computer*. 51, 9 (Sep. 2018), 28–36. DOI:https://doi.org/10.1109/MC.2018.3620965.

[10]   Kaye, J. "Jofish" et al. 2018. Panel: Voice Assistants, UX Design and Research. *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18* (Montreal QC, Canada, 2018), 1–5.

[11]   Kiseleva, J. et al. 2016. Understanding User Satisfaction with Intelligent Assistants. *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval - CHIIR '16* (Carrboro, North Carolina, USA, 2016), 121–130.

[12]   Lahoual, D. and Frejus, M. 2019. When Users Assist the Voice Assistants: From Supervision to Failure Resolution. *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow Scotland Uk, May 2019), 1–8.

[13]   Lau, J. et al. 2018. Alexa, Are You Listening?: Privacy Perceptions, Concerns and Privacy-seeking Behaviors with Smart Speakers. *Proceedings of the ACM on Human-Computer Interaction*. 2, CSCW (Nov. 2018), 1–31. DOI:https://doi.org/10.1145/3274371.

[14]   Lopatovska, I. and Williams, H. 2018. Personification of the Amazon Alexa: BFF or a Mindless Companion. *Proceedings of the 2018 Conference on Human Information Interaction&Retrieval - CHIIR '18* (New Brunswick, NJ, USA, 2018), 265–268.

[15]   Luger, E. and Sellen, A. 2016. "Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16* (Santa Clara, California, USA, 2016), 5286–5297.

[16]   Molnar, C. 2020. *Interpretable machine learning: a guide for making black box models explainable*.

[17]   Moore, R.J. et al. 2017. Conversational UX Design. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17* (New York, New York, USA, 2017), 492–497.

[18]   Nieva, R. and Robin, B.F. 2019. Amazon Alexa adds new commands to tamp down privacy concerns. *CNET*. (Sep. 2019).

[19]   Pins, D. et al. 2020. "Miss Understandable": a study on how users appropriate voice assistants and deal with misunderstandings. *MuC '20: Proceedings of the Conference on Mensch und Computer* (Sep. 2020), 349–359.

[20]   Purington, A. et al. 2017. Alexa is my new BFF. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17* (New York, New York, USA, 2017), 2853–2859.

[21]   Reeves, S. et al. 2018. Voice-based Conversational UX Studies and Design. *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18* (Montreal QC, Canada, 2018), 1–8.

[22]   Schönherr, L. et al. 2020. Unacceptable, where is my privacy? Exploring Accidental Triggers of Smart Speakers. *arXiv:2008.00508 [cs]*. (Aug. 2020).

[23]   Sciuto, A. et al. 2018. "Hey Alexa, What's Up?": A Mixed-Methods Studies of In-Home Conversational Agent Usage. *Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS '18* (Hong Kong, China, 2018), 857–868.

[24]   Shannon, C.E. and Weaver, W. 1998. *The mathematical theory of communication*. Univ. of Illinois Press.

[25]   Voit, A. et al. 2020. 'It's not a romantic relationship': Stories of Adoption and Abandonment of Smart Speakers at Home. *19th International Conference on Mobile and Ubiquitous Multimedia* (Essen Germany, Nov. 2020), 71–82.