Acceptance by Design: Voice Assistants

MARIA RAUSCHENBERGER, Faculty of Technology, University of Applied Sciences Emden/Leer, Germany

CCS Concepts: • **Human-centered computing → Interaction design**; *Human computer interaction (HCI)*; *Ubiquitous and mobile computing*.

Additional Key Words and Phrases: Design process, Speech-based Technology, Technology Adoption, Voice User Interface, VUI, Voice Assistant, User Experience, UX, Human-Centered Design Process, Context of Use

1 INTRODUCTION

Voice Assistants (VAs), Speech-based Technology (SBT), and Voice User Interfaces (VUI) have gained popularity in recent years; however, privacy concerns stop people from using them. Although the terms are used for slightly different purposes, we mainly use "VAs" for simplicity. To increase the adoption rate and address users' needs, we propose reducing barriers with *Acceptance by Design*. By that, we mean using, *e.g.,* VA designs that implicitly show if they collect data or how much data is collected. We can do that by first exploring the user requirements and concerns with methods from the Human-Centered Design Process [4] to provide guidelines for the future. This is a complex challenge that requires a deep understanding from different disciplines to design VAs for users skeptical about them. The HCD offers a holistic approach for capturing user requirements or discovering new dimensions of the UX that influence the user's perception or awareness of the interactive system. We must keep in mind that users have different needs (*e.g.,* children *vs.* parents *vs.* people with disabilities) that could result in different design solutions. Research has already started to explore with questionnaires [9, 10] and interviews the risks and opportunities of VAs (*e.g.,* cultures [6, 22], use cases [1, 3], existing users, or potential users [1, 3, 22]).

Furthermore, these guidelines might help others to design future interactive systems for different use cases, such as augmented reality glasses for university students in labs.

2 WHAT DO I KNOW ALREADY?

Voice assistants (VAs) have become ubiquitous in recent years. With the rise of products such as Alexa (Amazon) and Siri (Apple), VA use is expected to increase even further [23]. Not only do VAs have many advantages, but users already use speech commands while driving or changing music in a car. Future use cases might include commands while doing something else, such as cooking. Other opportunities depend on the user's abilities, such as for children who cannot write yet or people with disabilities, *e.g.,* visual impairments. While parents often have privacy concerns, children tend to be less fearful using new systems and therefore take advantage of a voice interface. People with disabilities use VAs intensively, not only for goal-orientated tasks, but also for pleasure, *e.g.,* games [8].

In general, users from Germany have shown high VA usage, but they also have concerns about privacy due to the collection of data [3, 7, 22]. This fact comes as no surprise, since VAs are placed in private spaces that are highly sensitive. Information and conversations originally did not leave spaces such as living rooms, bathrooms, or bedrooms. With VA microphones that are triggered by commands such as "*Alexa*," users feel like there is a fly on the wall that is always listening.

3 HOW DO I STUDY THE PHENOMENON?

Innovations trigger concerns and therefore require early adopters to explore the use cases and lower the barriers for using new technologies such as VAs. But more importantly, how we design such systems (*e.g., VAs, SBTs, VUIs*) with the potential to invade private spaces could have a great impact on the adoption rate of users and lower the barrier for using such systems. Therefore, we propose designing VAs with the *Human-Centered Design Process* (*HCD*) [4] to address users' requirements and concerns. However, it is not enough to just design with the HCD to overcome users' concerns and have them use the benefits of speech-based technology for their daily routine and different use cases independent of their ability. We also need to understand the triggers and use cases of users' concerns in order to derive design guidelines or patterns to prompt *Acceptance by Design*.

We have started to explore user risks and potential with technology-based users, as this target group shows a range of different interactions as well as privacy concerns [7].

When that is finished, we can derive user requirements not only for the usage, but also for the design of the Voice User Interfaces. To the best of our knowledge, there are no guidelines like those for Graphical Interfaces, such as the "Nielsen Heuristic" or "Shneiderman's 8 golden rules" [2]. The same is true for the design of the system and the evaluation of VAs.

Proposals for measuring the VA user experience exist, but have not yet been evaluated [5, 10, 10].

Hence, we need to explore the contexts of use and derive guidelines to design VAs that are *accepted by design*. We must also develop tools to evaluate our VA in order to improve the adoption rate through the affordance of the design. An example could be using ambient light to encode whether or not the microphone is listening [11]. In addition, we can design prototypes with different users, design concepts, and use cases and employ qualitative (*e.g.,* interviews) and quantitative evaluation (*e.g.,* questionnaires for VUIs [9]) to determine which designs are accepted by users or raise concerns. I also developed systems for user groups that need more protection, such as children [16, 19, 20]. In such cases, not only the user (a child) needs to be considered, but also the legal guardians (*e.g.,* parents) need to be included in the design process. Otherwise, the new system may not be successful if parents do not allow their children to use it.

In other research projects, I focused on how to handle small data in machine learning (prediction and clustering) [14, 15, 24]. This is the first step towards valid and solid results that can be used to make design choices. Next, we could explore ways to use small data analysis to improve speech recognition and response quality.

4 WHAT WOULD I LIKE TO KNOW?

When the aim is to design a system that focuses on the adoption of a new technology or its acceptance through design, the questions asked can include: "*How can one design VAs so that they provide privacy or transparency through their design?*" and "*Can we ensure privacy with our design choices, as is done for light encoding [11]?*"

Attributes could focus on how much data is collected or if the microphone is "*listening*". The main difference between public cameras and technology in a user's home, such as VAs, is the invasion of personal space. Therefore, figuring out how to design a VA system that is accepted and, for this example, focuses on privacy concerns, is also important for any other new technology that captures data from individuals (such as cameras, augmented reality technology, or data glasses).

The main problem seems to be that users do not know if their private conversations are being captured, leading them to ask questions such as: "*How much is captured?*" and "*What is happening with this captured data*?" The lack of transparency around sensitive data handling makes users uncomfortable. At the same time, the data is needed to improve speech recognition and response quality [13].

Hence, there are three possible solutions: propose new ways to increase speech recognition or response quality, collect less data, or design the system to increase transparency. This is a complex and multidisciplinary challenge that requires expertise from different backgrounds, *e.g,* machine learning for small data prediction, design knowledge, and user studies. When these approaches are explored and user needs are understood, the adoption rate for VAs might increase and privacy concerns might decrease.

5 WHAT DO I WANT TO LEARN FROM DIFFERENT DISCIPLINES?

Exploring and designing VAs can benefit from both the well-known methods, artifacts, and evaluation approaches in the domain of the HCD [2, 4] as well as the new approaches of small data analysis in machine learning (ML) [14, 15, 24]. The goal should be to increase the acceptance of the new technology by its design. We must think about further ways to adapt existing methods and develop new methods or artifacts from other disciplines which might not yet be known.

For example, we can use existing methods to collect user requirements and cluster them with the context analysis [12]. The advantage of the context analysis is the structured way in which assumptions are backed up or put into context by users' statements. Additionally, we should consider using the existing UEQplus Framework with newly developed VUI scales to measure VAs [9, 21].

Another way might be to avoid collecting big data with microphones and instead use small data for the machine learning analysis to increase the quality of speech recognition and responses, as is necessary in other domains due to the lack of data [14, 15, 24]. From previous research, we can learn how to use small data for prediction in the VA domain.

To sum up, I would like to know about methods from other disciplines that could be integrated into the design approach of VAs to create better systems and to cluster the requirements of good systems.

6 WHAT DO I WANT TO TEACH OTHER DISCIPLINES?

I want to teach other disciplines how to include user contexts and requirements to design good systems. We are already exploring these areas to better understand VA users' needs. Technology-based users from Germany have access to VAs, but do not use them because of privacy concerns. Hence, users' needs and concerns should be considered to design better systems, thereby reducing concern and increasing the usage rate.

As a result, evaluation tools such as the UEQplus [21] will include VA scale dimensions [9] to support the design process of VAs. This is just one example of how HCD and its methods can help to lower the barriers of new technologies such as VAs and increase their adoption rates.

I also want to raise awareness about how to handle small data analysis in order to prompt other disciplines to use even small data in machine learning [14, 15, 24]. This is because collecting less data might help users to trust more in the system. More controlled data and less data overall could mean, for example, that the data is collected in an initiation phase to improve speech recognition (e.g., due to a dialect). For this example, knowledge from linguists might be helpful to improve the machine learning algorithm, as we have already done in other projects [17, 18].

To sump up, I want to teach other disciplines what can be done already to design better systems and lower the concerns with the design.

REFERENCES

[1] Maresa Biermann, Evelyn Schweiger, and Martin Jentsch. 2019. Talking to Stupid?!? Improving Voice User Interfaces. https://doi.org/10.18420/muc2019-up-0253

[2] Henning Brau and Florian Sarodnick. 2006. *Methoden der Usability Evaluation (Methods of Usability Evaluation)* (2 ed.). Verlag Hans Huber, Bern. 251 pages. http://d-nb.info/1003981860http://www.amazon.com/Methoden-Usability-Evaluation-Henning-Brau/dp/3456842007

[3] BVDW e.V. 2017. Digital Trends Umfrage zu digitalen Sprachassistenten. Bundesverband Digitale Wirtschaft (BVDW) e.V. [Digital Trends Survey on digital language assistants. Federal Association of Digital Economy]. https://www.bvdw.org/themen/publikationen/detail/artikel/digital-trends-umfrage-zu-digitalen-sprachassistenten/. https://www.bvdw.org/themen/publikationen/detail/artikel/digital-trends-umfrage-zu-digitalen-sprachassistenten/

[4] DIN Deutsches Institut für Normung e. V. 2020. *DIN EN ISO 9241-210:2020-03, Ergonomie der Mensch-System-Interaktion - Teil 210: Menschzentrierte Gestaltung interaktiver Systeme; Deutsche Fassung.* Technical Report. Beuth Verlag GmbH. https://doi.org/10.31030/3104744

[5] Kate Hone. 2014. Usability measurement for speech systems : SASSI revisited. "https://www.semanticscholar.org/paper/Usability-measurement-for-speech-systems-%3A-SASSI-Hone/5db24db011fd5b867b95ac29b0b1085dc552eef9"

[6] Bret Kinsella and Ava Mutchler. 2018. Voice Assistant Consumer Adoption Report. https://voicebot.ai/wp-content/uploads/2019/01/voice-assistant-consumer-adoption-report-2018-voicebot.pdf. https://voicebot.ai/wp-content/uploads/2019/01/voice-assistant-consumer-adoption-report-2018-voicebot.pdf

[7] Andreas Klein, Andreas Hinderks, Maria Rauschenberger, and Jörg Thomaschewski. 2020. Exploring Voice Assistant Risks and Potential with Technology-based Users. , 147-154 pages. https://doi.org/10.5220/0010150101470154

[8] Andreas M. Klein. 2021. Toward a User Experience Tool Selector for Voice User Interfaces [DC]. *W4A'21, April 19-20, 2021, Ljubljana, Slovenia* (2021), 2–3. https://doi.org/10.2196/18431.4

[9] Andreas M. Klein, Andreas Hinderks, Martin Schrepp, and Jörg Thomaschewski. 2020. Construction of UEQ+ scales for voice quality. In *Proceedings of the Conference on Mensch und Computer*. ACM, New York, NY, USA, 1–5. https://doi.org/10.1145/3404983.3410003

[10] Andreas M. Klein, Andreas Hinderks, Martin Schrepp, and Jörg Thomaschewski. 2020. Measuring User Experience Quality of Voice Assistants. In *2020 15th Iberian Conference on Information Systems and Technologies (CISTI)*. 1–4. https://doi.org/10.23919/CISTI49556.2020.9140966

[11] Andrii Matviienko, Maria Rauschenberger, Vanessa Cobus, Janko Timmermann, Heiko Müller, Jutta Fortmann, Andreas Löcken, Christoph Trappe, Wilko Heuten, and Susanne Boll. 2015. Deriving design guidelines for ambient light systems. In *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia*, Vol. 30-Novembe. ACM, New York, NY, USA, 267–277. https://doi.org/10.1145/2836041.2836069

[12] Philipp Mayring. 2000. Qualitative Content Analysis. http://www.qualitative-research.net/index.php/fqs/article/view/1089/2385

[13] Martin Porcheron, Joel E. Fischer, and Sarah Sharples. 2017. "Do Animals Have Accents?": Talking with Agents in Multi-Party Conversation. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) *(CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 207–219. https://doi.org/10.1145/2998181.2998298

[14] Maria Rauschenberger and Ricardo Baeza-Yates. 2020. How to Handle Health-Related Small Imbalanced Data in Machine Learning? *i-com* 19, 3 (2020), 215–226. https://doi.org/10.1515/icom-2020-0018

[15] Maria Rauschenberger and Ricardo Baeza-Yates. 2020. Recommendations to Handle Health-related Small Imbalanced Data in Machine Learning. In *Mensch und Computer 2020 - Workshopband (Human and Computer 2020 - Workshop proceedings)*, Bernhard Hansen, Christian AND Nürnberger, Andreas AND Preim (Ed.). Gesellschaft für Informatik e.V., Bonn, 1–7. https://doi.org/10.18420/muc2020-ws111-333

[16] Maria Rauschenberger, Ricardo Baeza-Yates, and Luz Rello. 2020. Screening Risk of Dyslexia through a Web-Game using Language-Independent Content and Machine Learning. In *W4a'2020*. ACM Press, Taipei, 1–12. https://doi.org/10.1145/3371300.3383342

[17] Maria Rauschenberger, Silke Füchsel, Luz Rello, Clara Bayarri, Jörg Thomaschewski, F Silke, and Luz Rello. 2015. Exercises for German-Speaking Children with Dyslexia. *Human-Computer Interaction–INTERACT 2015* 9296 (2015), 445–452. https://doi.org/10.1007/978-3-319-22701-6

[18] Maria Rauschenberger, Silke Füchsel, Luz Rello, and Jörg Thomaschewski. 2016. A Language Resource of German Errors Written by Children with Dyslexia.. In *The International Conference on Language Resources and Evaluation — LREC 2016*. European Language Resources Association (ELRA), Portorož, Slovenia, 83–87. http://www.lrec-conf.org/proceedings/lrec2016/summaries/136.htmlhttps://repositori.upf.edu/handle/10230/32454

[19] Maria Rauschenberger, Christian Lins, Noelle Rousselle, Sebastian Fudickar, and Andreas Hain. 2019. A Tablet Puzzle to Target Dyslexia Screening in Pre-Readers. In *Proceedings of the 5th EAI International Conference on Smart Objects and Technologies for Social Good - GOODTECHS*. Valencia, 155–159.

[20] Maria Rauschenberger, Luz Rello, and Ricardo Baeza-Yates. 2019. Technologies for Dyslexia. In *Web Accessibility Book* (2 ed.), Yeliz Yesilada and Simon Harper (Eds.). Vol. 1. Springer-Verlag London, London, 603–627. https://doi.org/10.1007/978-1-4471-7440-0

[21] Martin Schrepp and Jörg Thomaschewski. 2019. Design and Validation of a Framework for the Creation of User Experience Questionnaires. *International Journal of Interactive Multimedia and Artificial Intelligence* (2019), S. 88–95. https://doi.org/10.9781/ijimai.2019.06.006

[22] Serpil Tas, Christian Hildebrandt, and René Arnold. 2019. Voice Assistants in Germany. https://www.wik.org.

[23] Sven Tuzovic and Stefanie Paluch. 2018. *Conversational Commerce – A New Era for Service Business Development?* Springer Fachmedien Wiesbaden, Wiesbaden, 81–100. https://doi.org/10.1007/978-3-658-22426-4_4

[24] Anna Christina Weigand, Daniel Lange, and Maria Rauschenberger. 2021. How can Small Data Sets be Clustered ?. In *Mensch und Computer 2021}{Workshopband}{Workshop on User-Centered Artificial Intelligence (UCAI '21)*. Association for Computing Machinery. https://doi.org/10.18420/muc2021-mci-ws02-284