Data Capture and Analyses from Conversational Devices in the Homes of the Elderly

Sandeep Purao, Chenhang Meng
Bentley University, Waltham, MA, USA
Email: spurao@bentley.edu

Abstract. Conversational devices such as Amazon Echo and Google Home represent more than a way to tap into the behavioral surplus of consumers. They provide an opportunity to address societal problems by examining data streams produced by these devices. In this paper, we describe usage patterns and problems related to the use of Amazon Echo devices at home by one specific demographic: the elderly. We rely on a pilot project to collect usage data over multiple months based on deployment of these devices in the homes of eight elderly individuals who either live alone or with a spouse. The paper describes methods used to ensure confidentiality, data collection and analysis procedures, and our findings. We find that the use of conversational devices remains restricted to single commands instead of conversations, making yourself understood remains a problem, sustained use remains a challenge, and the interaction rarely goes beyond simple commands. We interpret the results, and point to the potential for such devices in the lives of the elderly, specifically for health-related problems. The paper also describes lessons learned for capture and analysis of data from such conversational devices.

Keywords: conversational devices, elderly, use patterns

1 Introduction

Conversational devices have mushroomed. Although exact numbers are difficult to pin down [Lopatovska et al. 2018], it is estimated that more than 25 million people today use devices such as Amazon Echo™, Google Home™ and others[1] [Reis et al. 2017]. In spite of their almost ubiquitous presence of these devices, much of what we know about the adoption and use of these devices comes from industry reports. For Amazon and Google, these devices represent the possibility of "colonizing" everyday space such as home and office [Atlantic 2018], by connecting to an ecosystem of other devices and services. Amazon Echo already works with more than 20,000 smart-home devices representing more than 3,500 brands [Atlantic 2018]. It has been reasoned that this shift is likely to be wide and profound with the possibility of anthropomorphizing, bringing these devices closer to our own level. According to some accounts, "we [would] communicate with them, not through them" [Atlantic 2018]. Their roles today,

[1] The home-based conversational devices (the focus of this paper) are different from those embedded in mobile devices (e.g. Cortana™, and Siri™).

as devices that deliver news, calendar, weather, recipes, reminders and others are harbingers of what is to come. It is, therefore, important for us to better understand our interactions with these devices. Most of what we know about the use of these devices comes from industry reports [Ong and Suplizio 2016]. Few studies have examined the use of such devices (e.g. [Sciuto et al. 2018]).

In this study, our intent is to focus on the use of a dominant example of conversational devices, Amazon Echo, by a specific segment of the population, the elderly. This is an important population segment for multiple reasons. First, the elderly are a growing segment [Giacardi et al. 2016] that present unique challenges for technology design and use. It is, therefore, important to understand how the elderly use these conversational devices. Second, aging can lead to limited mobility, visual and hearing impairments, and high illness susceptibility [Khoury et al. 2018]. These characteristics define the values that these older citizens hold dear, often different from their younger counterparts as seen elsewhere in the design of service platforms [Skouby et al. 2014; Gil and Amaro 2010]. The conversational devices represent a timely example that we can investigate to better understand how the promise of technology use by the elderly may be realized.

The primary goal of this work is, therefore, to take first steps towards understanding how the elderly use such conversational devices, and the problems they face. A second goal for this work is to demonstrate how the large streams of data from these conversational devices may be captured and analyzed. We respond to these goals by carrying out a research project that captures and analyzes data from home-based conversational devices deployed in the homes of the elderly, while maintaining important privacy and confidentiality. Key contributions of our work include: findings related to how the elderly use conversational devices, and lessons related to capture and analysis of data from such conversational devices while maintaining privacy and confidentiality. The reminder of the paper is organized as follows. Section 2 reviews prior work. Section 3 describes the research setting and the research approach. Section 4 discusses the findings. In section 5, we wrap up with some concluding remarks.

2 Background and Prior Work

The population on the planet is aging – growth rates for the elderly are twice that of the overall population [Beard et al. 2016]. The skyrocketing costs of healthcare and services for the elderly are indications of these trends [Iwasaki 2013]. The design of appropriate information technologies is one key response to these trends.

2.1 The Design and Use of Technology for and by the Elderly

The phrase ‘elderly’ or ‘aging’ describes individuals who are 65 and over. However, with advances in healthcare, better diet and exercise, the stereotype of the elderly as ‘frail, vulnerable, immobile and passive’ is being questioned [Harvey and Thurnwald 2009, Kendig and Browning 2011]. Instead, there is a much larger group of healthy, active, independent “young old” who possess a very different image of themselves

[Giacardi et al. 2016]. Today's elderly may have spent a better part of their working lives learning new technologies. Therefore, some of the traditional lessons about "gerontechnology" [Kwon et al. 2016] are being challenged. There is a greater recognition that there is great variety in the everyday lives, needs and motivations of the elderly; and technology design is moving away from making things that are "foolproof" [Hyysalo 2006] to making technology that leads to 'more resourceful aging' [Giacardi et al. 2016]. This move to a more enlightened view describes technology design for the elderly in a manner that is ethical and responds to the concerns and values of the elderly [Giacardi et al. 2016; Purao et al. 2015]. It is in this context that we study the conversational agents at home for the elderly. They are not necessarily foolproof devices. Rather, they represent a set of capabilities that would allow the elderly to participate in resourceful aging, while still acknowledging that the elderly may not be fully prepared to use all new technological tools.

2.2 Conversational Agents at Home

As pointed out earlier, these conversational agents come in many forms, embedded in our mobile devices as well as location-locked devices such as Google Home and Amazon Echo. They respond to commands that follow what is referred to as a wake-word, e.g. 'Alexa' in case of Amazon Echo. A typical interaction with such a device starts with a command such as "Alexa, What's the weather today?" that results in a response such as "In Boston, it is 80 degrees with a 20% chance of rain." The devices contain software that captures the voice commands, converts these to text, develops a response, which may require accessing resources on the internet and speaks the response. Specialized capabilities can be developed for these devices (e.g. Skills for the Amazon Echo), which allow the device to respond to more specialized commands. According to some analyses, these devices represent an entry into the private homes of individuals with the possibility of tapping into so-called behavioral surplus [Zuboff 2019]. Others point to the potential of such devices to become a normal, persistent and important part of many households [Scuito 2018]. Studies that have examined the use of these devices have provided description of use such as the placement of devices at home, types of commands used and other similar statistics [Sciuto et al. 2018]. Such studies have remained rare. In this study, we hope to add to this nascent stream.

3 Research Approach

3.1 Research Setting

Driven by our research goal, we collaborated with the Council on Aging in one of the cities surrounding greater Boston. The city has a nominal population of ~60,000, with as one sixth of the population, 10,000+ are elderly. The city is home to large multiple corporate headquarters as well as universities. As a result, people who work in or travel to the city, are known to more than quadruple during the day. The Council on Aging is responsible coordinating activities and support services for the elderly with programs

such as meals on wheels, classes for the elderly, trips to different local facilities, visits to groceries and hospitals, and others. It occupies a separate facility in downtown. The Director of the Council participates in city administration along with the Mayor, District Councilors, and appointed members of the city government.

3.2 Subjects and Procedure

Working with this Council on Aging, we recruited individuals who expressed willingness to participate in the study. The research team specified these qualifications for participation: (a) have Wi-Fi available at home to use the Amazon Echo device, and (b) not presently have Amazon Echo or other conversational device at home. Following appropriate research protections, a sign-up sheet was made available at the Council. Individuals who expressed willingness were contacted by the research team. To deploy the devices, the research team visited the homes and activated the devices.

This process did not use any of the personal information of the individuals. Instead, the research team generated Amazon accounts and corresponding Gmail accounts as dummy identifiers. The dummy identifiers ensured privacy, but also limited use of the Amazon Echo devices for purposes such as shopping, ordering transportation and others, controlling other devices at home or making personal phone calls. By removing these, we were then able to focus on the use of Amazon Echo as a truly conversational device, instead of using it as, for example, 'a shopping assistant,' or 'a home hub' or other such modes. Minimal training was provided. A one-page document allayed fears about how to use the device (e.g. you can say 'Alexa, stop' any time), and showed some possibilities such as Alexa, what's the weather today; Alexa, should I carry an umbrella; Alexa, is the pharmacy open today; and so on. The research team retained access to the credentials, which allowed scraping of data about how each individual used the device. A mapping to codes with different individuals responsible for (a) deploying the devices, and (b) scraping the data ensured privacy during this process.

3.3 Data Scraping and Analyses

To access the history of Amazon Echo use, we scraped data for each user by using the normal login process. Figure 1 shows an example from one of the participants.

8	7/18/19	5:34	PM	alexa	117	16	Thursday
9	7/18/19	5:26	PM	how much ti	117	16	Thursday
10	7/18/19	5:11	PM	set a timer f	117	16	Thursday

Fig. 1. Data Scraped for a Specific User from Amazon Alexa: An Example

A Python script read through the front-end and dealt with problems of differential display (text, time), and partial display (screen size, week). Privacy safeguards remained. To facilitate analysis, additional data was inferred such as days the device has been active, day of week, time, an indicator to capture whether the command was understood, and time elapsed since the previous interaction. The data consisted of 7,829 commands from 8 users over 6 months. Following the exploratory nature of the study, we allowed the data to guide our analyses. We describe the findings next.

4 Findings

We begin with a simple conceptual model of the data recorded and generated by the Amazon Echo device (see Figure 2). To ensure confidentiality, not all the data elements (e.g. recorded user voice commands) were captured.

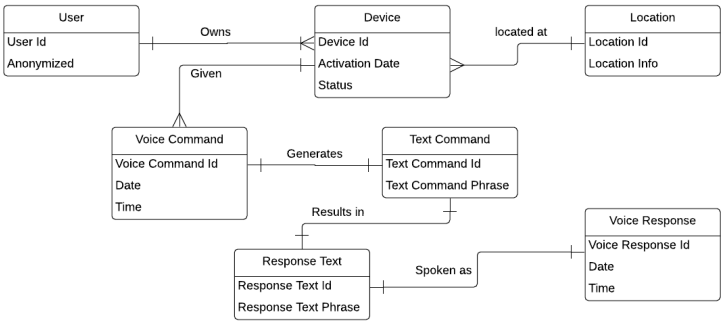
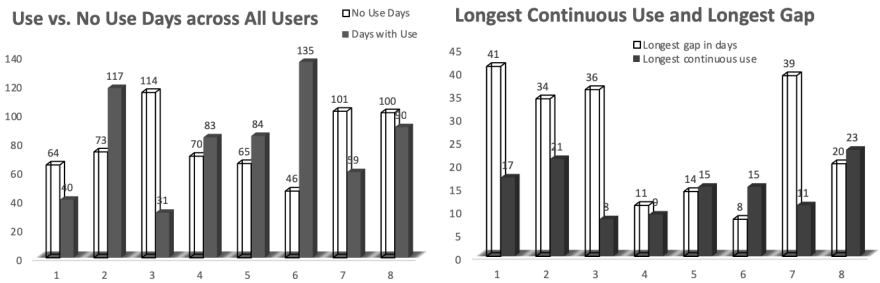


Fig. 2. Conceptual Model of Data Scraped from Alexa

Depending upon the scale, the conceptual model may be translated to a star-schema with the facts of interest as the core, to make the analyses more efficient. The model was populated with the data captured. The analysis results follow next.

4.1 Sustained Use and Intensity of Use

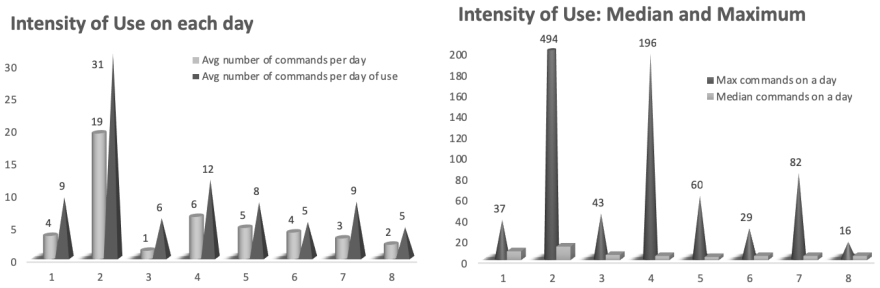
A prerequisite to conceptualizing these devices as intelligent personal assistants [Reis et al. 2017] is frequent and sustained use. The first set of analyses, therefore, examines analyzes how often and how regularly the elderly individuals use these devices: number of days the device was used vs. not used; and the longest contiguous number of days the device was used vs. the longest gap between uses. Figure 3 shows these results. The X-axis indicates the users.



3A. Days with Use vs. No Use Days 3B. Continuous Use vs. Gaps

Fig. 3. Sustained Use across All Users

The data shows that most users were active at least for some of the days. Figure 3A shows that the number of days with use was higher than the number of non-use days for users 2, 4, 5 and 6. Figure 3B shows that users 5, 6, and 8 had more contiguous days of use compared to the gaps with user 6 showing the most consistent usage with the largest gap between uses as 8 days, compared to users 1, 2, 3 and 7 who had gaps of 30 days or more. Together, the two graphs show that users 3 and 7 were not enthusiastic users of the device. The remaining users suggest possible sustained use. The next set of analyses examined use intensity: average number of commands (per day vs. per day of use); and median vs. maximum commands per day of use (see Figure 4).



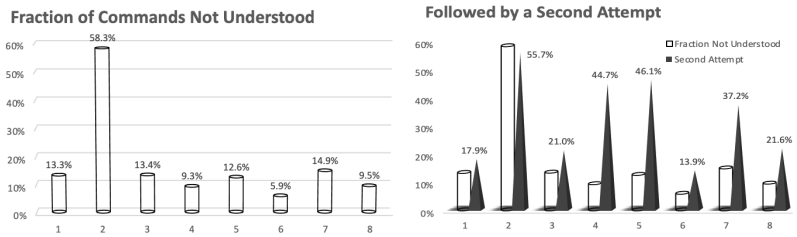
4A. Number of Commands / Day 4B. Median and Max Commands / Day

Fig. 4. Intensity of Use across All Users

The results continue to show users 2 and 4 as the most intense users of the devices. For example, user 2 used more than 30 commands on average on days that s/he was using the device with one of the days showing 494 commands (a clear outlier). User 4 used more than 12 commands on average on days that s/he was using the device with one of the days showing 196 commands (another outlier). Interestingly, user 7 showed a spike on one day with 82 commands (see Figure 4B), and the his/her average number of commands on the days of use was 9 (only a little behind users 2 and 4). Based on the data in figure 4A, users 1 and 7 were next to users 2 and 4 in terms of average number of commands per day of use in spite of long gaps in usage (see Figure 3B).

4.2 Commands: Some Understood, Others Not So

To understand whether this use actually resulted in responses, we examined whether the commands of the elderly individuals were actually understood by the conversational devices. This is indicated by the classification 'text not understood' in the data scraped. Such a response from the device meant that when the elderly individuals attempted to use the devices, their command was not understood. While this may not be as significant a concern for other sets of users, it can be significant for the elderly (e.g. see [Kwon et al. 2016]). When this occurred, in some cases, the individuals appeared to repeat the command (either with the same words or different). Figure 5 shows the results of our analysis. The x-axis shows the users, and the y-axis shows the fraction of commands not understood, followed by a second attempt (Figures 5A and 5B).



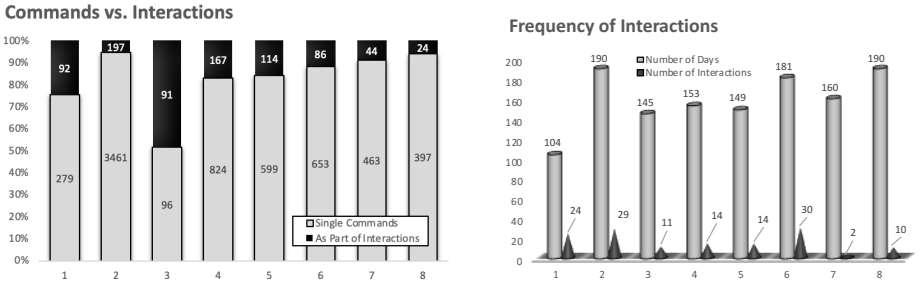
5A. Fraction Not Understood 5B. Followed by a Second Attempt

Fig. 5. Commands (Not) Understood as Intended, across All Users

The data shows that user 2 was a clear outlier with as many as 58% of commands resulting in 'text not understood,' with 55.7% also unsuccessful as second attempts. With this new information, the characterization of user 2 as a frequent user (see Figures 4A and 4B) now became suspect. Although this user attempted to use the device a lot, more than half of his/her commands were not understood. All the other users did reasonably well with less than 15% of the commands not understood (although these numbers may still be frustrating for some individuals). Other users that stood out in this analysis included users 4, 5 and 7 (see Figure 5B). These users used a second attempt to clarify their commands and were still unsuccessful 37 to 46% of the times. Together, these results show that using these devices still remains problematic for the elderly.

4.3 Commands, *Not* Interactions (or Conversations)

Next, we explored the central idea that these are 'conversational' devices [Atlantic 2018], i.e., more than a single command-response pair. We were, therefore, interested in examining the nature of conversations that the elderly individuals were having with these devices. Our simple conceptualization of a conversation was any interaction that was longer than a single command-response pair. To explore this, we separated commands that were part of a larger "interaction (or conversation) episode," against commands that were part of a single "command-response pair" (with a parameter value of 3 minutes to separate one episode from the next). We also compared the number of "interaction (or conversation) episodes" against the number of days (see Figure 6).



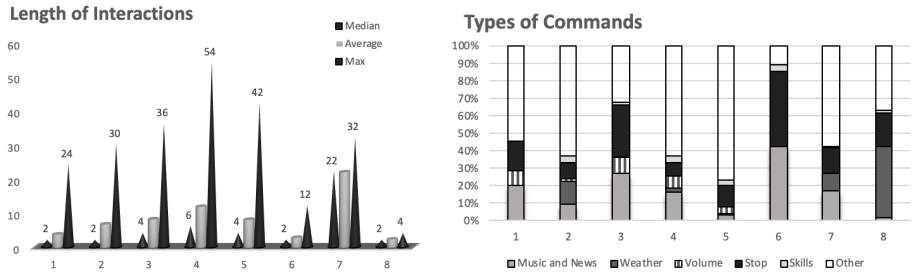
6A. Commands vs. "Interactions" 6B. Frequency of "Interactions"

Fig. 6. Commands, *Not* Interactions or Conversations, across All Users

The results show that only user 3 managed to have such interactions on a somewhat regular basis (see Figure 6A) with 91 of the commands part of such "interaction (or conversation) episodes" compared to 96 commands that were part of a simple command-response pair. Even with such a high fraction, user 3 managed only 11 such "episodes" over the entire duration of 145 days. User 1, in fact, had more such "episodes" (24) over a fewer days (104) with 92 commands as part of these "episodes" (see Figures 6A and 6B). User 2 managed only 2 interactions over the entire period of 160 days.

4.4 Characterizing Conversations and Commands

Finally, even with the problems with sustained use (see Figures 3 and 4), making yourself understood (see Figure 5), and engaging in simple command-response pairs instead of longer episodes (see Figure 6) – the elderly individuals still managed use these devices. To explore these, we considered the length of the interaction (or conversation) episodes, as the number of commands within each episode (Figure 7A), and examined the types of commands they used with categories similar to prior work [Scuito et al. 2018] (Figure 7B). Figure 7 shows these results.



7A. Length of Interactions/Conversations 7B. Types of Commands

Fig. 7. Characterizing Conversations and Commands, across All Users

The data showed that the average length of "interaction/conversation episodes" was fairly low for all users (see Figure 7A). The outlier (user 7) was an anomaly with an average interaction episode length of 22. However, this was misleading because it represented two episodes (one of length 12 and the other of length 32, with what appeared to be a set of unrelated commands). The results for user 4, on the other hand, had a median of 6 commands (see Figure 7A) across 14 episodes (see Figure 6B). Next, we examined the types of commands by examining different categories: music and news, weather, volume, stop, and skills. Figure 7B shows the percentage of commands in each. The users in our set had different emphases, further cementing the idea that the elderly cannot all be treated the same [Giacardi et al, 2016]). It is important to use one caveat with these results. No personal uses such as making family phone calls or shopping were part of our study because of our emphasis on keeping the results anonymized. Including these categories may change the outcomes.

5 Discussion and Concluding Remarks

In this paper, we have explored data streams from conversational devices (Amazon Echo) deployed in the homes of the elderly. It is important to note that our intent is to report exploratory analyses of usage patterns for these devices, not explore the strengths or weaknesses of specific devices. The paper described our approach to scraping, and analyzing the data with the use of automated scripts. The approach is scalable. In fact, during this research, we continued to work with the data stream as the data accumulated each week. The key contributions of our work include findings related to the use of conversational devices by the elderly, which include: intensity and sustained use of the devices, examining commands understood and not understood, characterizing interactions and conversations, and exploring different types of commands. The conceptual model we have created can also provide pointers for collecting and analyzing these data streams. The picture that emerged from the analyses shows that the elderly can have different profiles, and these differences are important to consider when designing voice skills for the elderly. It is important to consider the findings to better position the conversational devices as supports for resourceful aging. More specifically, our analyses also show that leveraging conversational devices will require overcoming basic obstacles such as more precise understanding of the commands, and representing and using context to support longer conversations.

Acknowledgements

We acknowledge participation from the elderly, and the Waltham Council on Aging. We also appreciate comments from the review team that have shaped the final version. The work reported has been funded by the National Science Foundation under award number 1641148. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).

REFERENCES

1. Atlantic, The. 2018. Is Alexa Dangerous? Oct 24. Online. *The Atlantic*. Accessed 30 July 2018. https://www.theatlantic.com/magazine/archive/2018/11/alexa-how-will-you-change-us/570844/
2. Beard, J.R., Officer, A.M. and Cassels, A.K., 2016. *The world report on ageing and health*. United Nations.
3. Bickmore, T.W., Trinh, H., Olafsson, S., O'Leary, T.K., Asadi, R., Rickles, N.M. and Cruz, R., 2018. Patient and consumer safety risks when using conversational assistants for medical information: an observational study of Siri, Alexa, and Google Assistant. *Journal of medical Internet research*, *20*(9), p.e11510.
4. Bjering, H., Curry, J. and Maeder, A.J., 2014. Gerontechnology: The importance of user participation in ICT development for older adults. In *HIC* (pp. 7-12).
5. Bouma, H., 2012. Foundations and goals of gerontechnology. *Gerontechnology*, *11*, pp.1-4.

6. Coyne, M., Thomas, C., Collimore, A., Franzese, C. and Hwang, C., 2017. Early user centered insights on voice integrated technologies through retrospective analysis. *Iproceedings*, *3*(1), p.e49.
7. Dall, T.M., Gallo, P.D., Chakrabarti, R., West, T., Semilla, A.P. and Storm, M.V., 2013. An aging population and growing disease burden will require alarge and specialized health care workforce by 2025. *Health affairs*, 32(11), pp.2013-2020.
8. Giaccardi, E., Kuijer, L., & Neven, L. (2016). Design for resourceful ageing : intervening in the ethics of gerontechnology. In P. Lloyd, & E. Bohemia (Eds.), Proceedings of DRS 2016, Design + Research + Society Future-Future-Focused Thinking: 50th Anniversary International Conference, Brighton, UK, 27-30 June 2016 (Vol. 1)
9. Gil, H. and Amaro, F., 2010. Active ageing and the role of ICT and assistive technologies: Reflections and discussion for their use in Portugal. *e-case & e-tech 2010*, pp.2750-2760.
10. Gutman, G., Kearns, W., Normie, L., Kort, H.S.M. and van den Berg, P.E.W., 2018. Addressing the needs of older adults through technology: the unique focus of the *International Society for Gerontechnology*.
11. Harvey, P. W., & Thurnwald, I. (2009). Ageing well, ageing productively: The essential contribution of Australia's ageing population to the social and economic prosperity of the nation. *Health Sociology Review*, 18(4): 379-386.
12. Karapanos, E. et al. 2009. User experience over time: an initial framework. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: 729–738. http://doi.org/10.1145/1518701.1518814
13. Kendig, H., & Browning, C. (2011). Directions for Ageing Well in a Healthy Australia. *Academy of the Social Sciences*, 31(2): 23-30.
14. Khoury, Y., Purao, S. and Duffy, M., 2018. The Influence of Values on the Use of Citizen Services: The Elderly Perspective. *ICIS* TREO.
15. Kwon, S. ed., 2016. *Gerontechnology: Research, practice, and principles in the field of technology and aging*. Springer Publishing Company.
16. Laranjo, L., Dunn, A.G., Tong, H.L., Kocaballi, A.B., Chen, J., Bashir, R., Surian, D., Gallego, B., Magrabi, F., Lau, A.Y. and Coiera, E., 2018. Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association*, *25*(9), pp.1248-1258.
17. Lopatovska, I., Rink, K., Knight, I., Raines, K., Cosenza, K., Williams, H., Sorsche, P., Hirsch, D., Li, Q. and Martinez, A., 2018. Talk to me: Exploring user interactions with the Amazon Alexa. *Journal of Librarianship and Information Science*, p.0961000618759414.
18. Ram, A., Prasad, R., Khatri, C., Venkatesh, A., Gabriel, R., Liu, Q., Nunn, J., Hedayatnia, B., Cheng, M., Nagar, A. and King, E., 2018. *Conversational AI: The science behind the alexa prize*. arXiv preprint arXiv:1801.03604.
19. Reis, A., Paulino, D., Paredes, H. and Barroso, J., 2017, July. Using intelligent personal assistants to strengthen the elderlies' social bonds. In *International Conference on Universal Access in Human-Computer Interaction* (pp. 593-602). Springer, Cham.
20. Sciuto, A., Saini, A., Forlizzi, J. and Hong, J.I., 2018, June. Hey Alexa, What's Up?: A mixed-methods studies of in-home conversational agent usage. In *Proceedings of the 2018 Designing Interactive Systems Conference* (pp. 857-868). ACM.
21. Shulevitz, J., 2018. Alexa, should we trust you. The Atlantic.
22. Skouby, K.E., Kivimäki, A., Haukiputo, L., Lynggaard, P. and Windekilde, I.M., 2014, May. Smart cities and the ageing population. In The 32nd Meeting of WWRF.
23. Zuboff, S., 2019. *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile Books.