PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Alexa, Siri and more: The impact of speech recognition on social behaviour and our responsibility as its creators

Andrea EBNER[1]

[1] Technichal University of Munich, Germany

ABSTRACT

Digitalization and globalization allow us to create machines that can be used all over the world. But we must consider what people in the target markets want. What do they think of the hidden microphones in our living rooms, why are they afraid of speech-control? Are people ready to give us their trust in our creation and their consent to share their privacy with us? Since the buzzword AI can be read in every newspaper and it seems like there is no other technical topic out there to talk about any more, I want to ask some questions. What the societies worries are and where may they come from and how can we handle our responsibility as the producer of change? The political debate is often ideologically. We as scientist and engineers can support the decision-makers and the society with giving the rational facts and provide simple, understandable explanations of what is going on in the tech world. Because it is not just about with which good or evil device we interact. Instead of being afraid of change it is a matter of HOW WE INTERACT and HOW WE TRANSFORM the world into a better one for all of us.

Keywords: Speech assistants, Society, Responsibility

1. INTRODUCTION

Science fiction authors have known it for a long time: speech assistants make our lives easier. They keep us up to date and inform us aboutnews or remind us of important dates. Built into a robot, they do unloved work or help us raising our children. They know everything about us and offer an entertainment experience adapted to our preferences. But what is the impact of human-like technology systems on social behavior? Furthermore, the role of an engineer and his or her responsibility regarding the evolving technologies is addressed.

2. DISCUSSION

2.1 Speech assistants available on the market

Since everyone can buy his or her own Smart Speaker (a speaker with an installed speech recording software and a connection to the internet to process the speech commands in the cloud), the technology of speech assistants has experienced a remarkable boom. Just a few years ago (since 2011), Siri on the iPhone was only able to create a simple calendar entry for you. Nowadays Alexa and Co. can report on the weather, switch the lights on and off in your Smart Home and order pizza for you.

Amazon has had a considerable influence on the upswing. The company has launched its own device with speech recognition and speech processing software which makes the shopping experience even more convenient and increases the acceptance of the assistants linked to company offers. Therefore, the company was criticized with regard to data protection and received the Big Brother Award in 2018 (1).

Nevertheless, Amazon is enjoying market success with the Alexa software in various devices. Other large providers followed and also brought their own smart speakers onto the market (before Google launched its speaker in Germany in 2017, the Google Assistant was only available in smartphones since 2012). Apple was the last one to introduce a speaker to the market; they offer a speakeronly in the US since 2018 . In order to compare the systems, there are various criteria

[1] andrea.ebner@tum.de

according to which the quality of the language assistants is evaluated. The Alexa software features the most natural speech, while the Google Assistant can answer questions in a mostly reliable and correct way (2).

2.2 How the assistants work

The functionality is simple: The software wakes up when a code word is spoken and then the instruction is recorded.The system sends the received command to the server of the software provider, where it is processed. In some cases, everything is recorded and then sent to the server for speech analysis to extract the commands and instructions. The result is finally returned, usually that is an information or an instruction for a device from a Smart Home. For example, the user will hear music (from the Smart Speaker) or the lights will start to dim or change color. The hardware consists, among other things, of (at least) a loudspeaker and a microphone, where in some cases, the latter can be switched off with a button, in order to avoid unwanted command recognition.

This functional simplicity makes the system attractive and comfortable for many users. According to the manufacturer (Amazon), more than 100 million devices on which the Alexa software is installed were sold by the end of 2018. This is the case for in-house products as well as high-quality loudspeaker systems, among other devices (3).

The attractiveness of Amazon's software may be further enhanced by the fact that private and commercial developers can extend the functionalities by programingso-called skills. One can compare these skills with the apps for smartphones. Thus, a large community can participate in the technology and establish a connection with the product world and the company.

2.3 Useage of speech assistants

During a study conducted by Capgemini Digital Transformation Institute in 2018, people in different countries were asked about their opinion and habits about speech assistants.While 13% of German interview participants aged over 18 already use stationary speech assistants, 27% can imagine doing so in the near future (4). Furthermore, 38% of those surveyed in the US, the UK, France and Germany wish the human assistants in call centers to be replaced with personalized speech assistants. They expect time savings and more convenience for themselves (5).

2.4 Privacy and data protection

The main aspects of criticism are concerning privacy and the related issue of data protection. Privacy is considered to be a shielded area in which one can develop oneself in a very personal way (6). The fact that the boundaries of this area are interpreted individually and differently can be seen in the way social media are being used. While some people post pictures of themselves in any situation in life, others delight their friends with photographs of food and landscapes, some others strictly refuse to participate in social networks from the very beginning (7).

Furthermore, this definition of the private sphere is culturally and historically shaped. In recent years, it was mainly the US, and thus especially Silicon Valley, and China that set the pace. They enabled the rapid development of technology and therefore the perspective on privacy has been influenced in different ways. What the United States and Germany (and Europe) have in common, however, is the basic approach: the individual is at the center of debates. In parts of Asia and Africa, on the other hand, the well-being of the community is more desirable than the satisfaction of the individual's needs (8).

Moreover, every user of a device with an incorporated speech assistant connected to the Internet has to trust the companies to a certain extent, because when purchasing and using the device, the provider will invade the private sphere to perform its services.

The critical aspect of data protection can basically be summarized in a few simple words: Users choose less privacy for more convenience.

2.5 Education

Speech assistantst seem to be a very popular support in education. Parents use assistants as playmates for their children, for example the talking doll called “My friend Cayla”. However, the contact of children with talking toys, which can answer simple questions and provide them with information, is a sensitive matter. For instance, children may not be aware of any danger when "Cayla" suddenly asks them to open the balcony door and let a stranger into the house. This would

have been made possible by an unprotected Bluetooth connection, which would have allowed anyone to connect directly to the doll. Therefore it was classified as a forbidden transmitter according to § 90 TKG (Telecommunications Act in Germany) (9).

When children grow up with talking, interactive dolls or other devices, they develop a certain behavior towardsthese objects and seemingly imagine them to be real playmates. Some manufacturers have reacted to reports of a harsh conversational tone due to speaking in command form (10) with a "child-friendly" mode, which requires "Please" and "Thank you" in the communication.

In order to protect children from over-consumption, there are also features to limit the usage time of speech assistants. Parents themselves also intervened to slowly and consciously teach the children that they have a robot in their hands and that it should not be equated with a human being. The explanation of the difference between humans and machines is essential in all future education.

2.6 Interpersonal and close relationships

Not yet explicitly researched, but nevertheless frequently discussed in the media is the influence on the perception of gender roles. Most voices of speech assistants are female. This gives the impression that women are servants and handlers and that they are only employed for supporting activities (11).

In some cases, relationship desires are projected onto speech assistants, creating an ideal image of a partner. Survey participants think of the assistant as a nanny, pet, surrogate mother, partner and/ or coach. This projection increases the expectations on potential, real life partners and technical devices can become a competition for humans. Nevertheless, the fear of addiction and of being dependent on technical assistants is also reported (12).

This raises the question of the value of a relationship: Would you rather interact with an ever-friendly personal assistant or would you invest your feelings and efforts into a person who will be there for you under all circumstances in life?

2.7 Guidelines for responsibility

Nowadays, experts seem to have little time to deal with moral and social issues. But the more complex the use of the technology is, the more requirements and concerns must be considered. In a guideline recently published by IEEE on the ethically correct handling of autonomous and intelligent systems (13), the scientific disciplines involved are approached and advised about how they should act in order to follow ethical and social conventions. The following aspects, among others, are taken from that guideline and are extended.

2.7.1 Be open and discuss with others

To be able to take ethical guidelines into account and successfully implement a technology, one must first be aware of the potential problems of new technologies. Due to the complexity of systems, it is urgent to seek help from as many other disciplines as possible. An interdisciplinary approach is essential when dealing with complex issues.

Equally fundamental for constructive debates is a mutual acceptance of substantially different views and an understanding of concerns from other areas. If you have doubts, it is essential to step out of your own shadow and ask other experts (and "normal" society) at an early stage.

2.7.2 Fear of technology?

At the moment it seems as if society is mainly afraid of new technologies (artificial intelligence, humanoid robots, ...). It is therefore advisable for experts to get involved in public discussions at an early stage. With rational facts and clear, understandable examples it is possible to deprive dystopias of their basis. Often, fears are unfounded and based on a lack of knowledge (Can there be an artificial intelligence if you don't know what intelligence actually is?).

It is the responsibility of experts to make sure that their achievements are used in a sustainable way and in the best interests of mankind.

3. CONCLUSIONS

There is no doubt, technolocigal progress has both positive and negative sides. While lots of users appreciate the help of digital assistants, some are worried about security, privacy and the future developments.

Communication has and will always be evolving. People will always communicate and interact with each other in some form, be it via text messages or via spoken word. The content matters, not the means of transferring them. Speech assistants will integrate themselves into family and social life, as many technical achievements (TV, PC, smartphone) have done before. Families or individuals decide for themselves how, for what purpose and to what extent they want to use certain devices.

Like everywhere else, there are skeptics and fanatics. Through matter-of-fact discussions, issues such as data protection can be resolved in a compromise that takes both the technical progress and individual needs into account. Those who are knowledgeable, i.e. technicians, researchers, teachers and others, have the responsibility towards society to educate the non-scientists. They have to explain how the technology works, how to keep control of it and how to eventually deactivate it. They, or to be more precise: we, must contribute to education and technology training. We have to discuss with those who are not aware of the possible consequences of new technological developments and we have to critically question our own achievements. We have to calm media and society down if the debate tends to go surrealistic; and we should speak out loud against fake news about our technological achievements.

Our challenge as experts should be: How can the strengths of people and machines be brought together without neglecting people?

REFERENCES

1. Big Brother Award Homepage. URL: https://bigbrotherawards.de/2018/verbraucherschutz-amazon-alexa.
2. Stiftung Warentest: Assistenten im Test. Test 4/2018; 2018. p. 34-37.
3. Bohn, D. Amazon says 100 million Alexa devices have been sold — what's next?. The Verge; 2019. URL: https://www.theverge.com/2019/1/4/18168565/amazon-alexa-devices-how-many-sold-number-100-million-dave-limp.
4. Vom Smartphone ins Wohnzimmer der Konsumenten: Digitale Sprachassistenten. Research Spotlight 2018-04; 2018. URL: https://www.bitkom-research.de/epages/63742557.sf/de_DE/?ObjectPath=/Shops/63742557/Categories/Presse/Spotlight/Research_Spotlight_201804_EU_Datenschutz_Grundverordnung[1].
5. Conversational Commerce. Capgemini Digital Transformation Institute; 2018. p. 11, 15.
6. German Basic Law. Art. 2 Abs. 1 i.V.m. Art. 1 Abs. 1 GG, 1949.
7. by-nc-nd/3.0/ Autor: Dr. rer.pol. Jan-Hinrik Schmidt für bpb.de. URL: http://www.bpb.de/gesellschaft/digitales/datenschutz/194364/praktiken-und-risiken-sozialer-medien.
8. by-nc-nd/4.0 Autor: Prof. Kai von Lewinski für bpb.de: Was Europa und die USA in Sachen Datenschutz unterscheidet. URL: http://www.bpb.de/gesellschaft/digitales/democracy/254254/datenschutz-in-der-eu-und-den-usa.
9. Hessel, S. "My friend Cayla" - eine nach § 90 TKG verbotene Sendeanlage?. JurPC-Web-Dok. 0013/2017. Abs. 1-39.
10. Rudgard, O. 'Alexa generation' may be learning bad manners from talking to digital assistants, report warns. The Telegraph; 2018. URL: https://www.telegraph.co.uk/news/2018/01/31/alexa-generation-could-learning-bad-manners-talking-digital/.
11. Genzmer, J, Wiese. T. Die gesellschaftliche Wirkung von weiblichen Sprach-Assistenten; 2019. URL: https://www.deutschlandfunkkultur.de/the-handmaids-fail-die-gesellschaftliche-wirkung-von.1264.de.html?dram:article_id=443058.
12. Bezaubernde Beziehungskiste–wie Alexa die geheimen Wünsche ihrer Nutzer erfüllt, Rheingold Institute; 2017. URL: https://www.rheingold-marktforschung.de/wp-content/uploads/2017/08/Presseinformation_Alexa_rheingold-Institut.pdf.
13. The IEEE Global Initiative, Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, First Edition. CC BY-NC-SA 4.0; 2019.