Position Talk: Ethically aligned voice user interfaces: Risks, Challenges and Opportunities

Stephan Sigg
Aalto University
stephan.sigg@aalto.fi

***Abstract*—This position talk discusses different position, opportunities and challenges posed by upcoming voice user interface platforms. In particular, the presenter will highlight privacy and legislation aspects and give insight into newly forming initiatives towards privacy-preserving speech processing withing the speech processing community. Exemplarily, the speaker will introduce ongoing efforts to address these identified challenges.**

***Index Terms*—Privacy and trust in voice user interfaces, Usable Security, Legislation, speech processing, Fuzzy cryptography**

I. Talk details

We witness the spread of voice user interfaces (VUIs), such as Siri, Alexa, or Cortana. As acoustical front-end for voice-based services, they take part in conversational social activities via a cornucopia of novel services and revolutionize customer interaction, human-device communication, as well as traditional human-to-human telecommunication applications. Furthermore, additional services exploit emotional state, health status, gender, age or identity. For all of these, recent advances in speech communication, such as synchronized recordings from distributed microphones, accurate speaker localization and estimation of speech direction are bound to further raise the usefulness of these services. The integration of artificial intelligence (AI) during the processing is able to establish scene understanding capabilities and evolves VUIs into conversational bots, able to participate actively in human interaction.

Significant part of the audio processing and analysis as well as interpretation and reasoning on its content are conducted in a remote cloud operated by the VUI manufacturer. Information about when information is recorded, what part of the data is shared, as well as the purpose or duration of storing the data and whether it is further used by other third parties remains intransparent. In addition, with other humans in the reach of the microphone, their speech and audio is also recorded and shared without explicit consent or even notification of the process. We have to come to realization that the novel services provided by VUIs are acquired at the expense of privacy and ethical principles. In a society penetrated by a massive amount of such VUIs, such ubiquitous privacy annihalition has the potential to change a society fundamentally.

Technological advance can not be halted and likewise to recent other services that have heralded the digital age, data is the currency also for VUIs. No personalized service is possible without sharing to some extend personal and private information. However, widespread acceptance of VUIs, and therby their success in private, public and industrial domains demands trust in the technology. Establishing of trust, though, is hardly possible in today's setting which largely cements hierarchies of big industrial players to maximize data gain at a minimal cost and with intentionally low transparency for the individual. This can be seen as an ethical problem because the technology challenges existing ethical principles. However, it can also be interpreted as a question of lesgislation, because VUIs have entered a domain that is not yet regulated in national or international law. On the other hand, the problem is also one in the domain of user interfaces, because proper user interface design for VUIs can achieve the needed transparency. But it is a question of AI and Data Science too, because AI provides the algorithms and solutions to potentially filter relevant and ethically conformative from privacy-violating content. Lastly, it is a question that falls in the domain of economics because the technology will only be widely adopted if it makes sense economically for major international players. The question of ethically aligned voice user interfaces requires a close interplay and collaboration between experts from various interdisciplinary fields.

II. Speaker Bio

Stephan Sigg is an Associate Professor at Aalto University in the Department of Communications and Networking. His research interests include the design, analysis and optimisation of algorithms for distributed and ubiquitous systems. Especially, his work covers proactive computing, distributed adaptive beamforming, context-based secure key generation and device-free passive activity recognition. Stephan is an editorial board member of the Elsevier Journal on Computer Communications and has been a guest editor for the Springer Personal and Ubiquitous Computing Systems Journal. He has served on the organizing and technical committees numerous prestigious conferences including IEEE PerCom, ACM Ubicomp and IEEE ICDCS. He is an associate editor of ACM IMWUT.