

Zukunft der Mensch-Maschine-Interaktion: Spracherkennung und -steuerung

Die Sprache ist eine zentrale Kulturtechnik und deren Rolle in der Entwicklung der menschlichen Zivilisation ist unumstritten. Gegenwärtig zeichnet sich der Trend ab, dass der Sprache eine immer größere Bedeutung in der Mensch-Maschine-Interaktion zukommt, da die Entwicklungen von Hard- und Software in den letzten Jahren große Fortschritte im Bereich der Spracherkennung gebracht haben. Diese Fortschritte ermöglichen ein breites Anwendungsspektrum für neue Interaktionsarten von Mensch und Maschine. So ist es beispielsweise den NutzerInnen von Smartphones bereits möglich, mittels der im Betriebssystem integrierten Sprache-zu-Text-Anwendungen, Nachrichten zu verfassen. Entsprechend besteht zudem die technische Möglichkeit, das Gesprochene in Echtzeit in andere Sprachen zu übersetzen. Weitere Anwendungsmöglichkeiten ergeben sich für Menschen mit eingeschränkter Mobilität. In diesem Anwendungsfeld spielt neben der Spracherkennung vor allem auch die Sprachsteuerung eine wesentliche Rolle. Für die NutzerInnen von Spracherkennungsanwendungen bedeutet das einerseits den Zugewinn an Komfort, da eine manuelle Interaktion mit der Maschine – sei es nun ein Smartphone, Tablet oder PC – nicht mehr notwendig ist, und andererseits einen Zugewinn an Autonomie.

Alle großen Technologieanbieter sind an der Weiterentwicklung von Spracherkennung und -steuerung beteiligt und bieten diese z. T. auch als eigene Geräte zum Verkauf an (z. B. Amazon Echo, Google Home etc.). Diese Geräte sind in die digitalen „Ökosysteme“ der jeweiligen Hersteller eingebettet, was vor allem aus Sicht der Datensouveränität und des Datenschutzes Fragen aufwirft. Außerdem haben viele Systeme technische Sicherheitslücken (Edu et al. 2020). Durch Verwendung eines bestimmten Assistenten über mehrere Geräte hinweg (Sprachsteuerungen gibt es mittlerweile vom Automobilbereich bis hin zu Haushaltsgeräten wie Waschmaschinen über Smart Home-Steuerungen und Fernsehgeräten), wird nicht nur das NutzerInnenprofil detaillierter, es bekommt der Sprachassistent auch immer mehr die Rolle eines Gatekeepers in der Kommunikation mit der Außenwelt.

Ein Whitepaper von Microsoft aus dem Jahr 2019 über Sprachtechnologie und digitale Assistenten widmet sich der Frage der Akzeptanz und den zukünftigen Potentialen dieser Technologie. Aus Sicht von Microsoft wird die „Touchscreen-Gesellschaft“ über kurz oder lang von der „Voice-Assistant-Gesellschaft“ abgelöst werden. Der Bericht umfasst auch die Ergebnisse einer Onlineumfrage. Sie zeigen, dass etwas über 40 % der Befragten dieser Technologie in Bezug auf ihre Privatsphäre kritisch gegenüberstehen. Die größten Sorgen wurden im Hinblick auf die Sicherheit der personenbezogenen Daten bzw. auf die potentielle Gefahr des latenten

Mithörens durch permanent aktive Mikrofone geäußert.¹ Im Frühjahr 2019 wurde in den Medien darüber berichtet, dass Amazon, Google, Apple und Microsoft die Aufnahmen von und Befehle an Sprachassistenten durch MitarbeiterInnen anhören, transkribieren und überprüfen ließen. Das Ziel davon sei es, die Spracherkennung zu verbessern.² Diese Praktiken werfen zahlreiche datenschutzrechtliche Fragen auf. Datenschutzfreundliche Einstellungen bieten derzeit keine der Anbieter an, es werden detaillierte Persönlichkeitsprofile erstellt und direkt mit konkretem Verhalten der AnwenderInnen verknüpft (Schaber et al. 2019). Der Zugewinn an Komfort geht somit mit einer Einbuße an Privatsphäre einher (Ferdinand/Jetzke 2017). Darüber hinaus scheinen sich neue Abhängigkeiten von großen US-Konzernen und chinesischen Firmen zu ergeben, die potentiell geopolitische Folgen haben können, da europäische Firmen derzeit sowohl bei den Sprachmodellen als auch der Verfügbarkeit von Trainingsdaten im Hintertreffen sind (Peters 2022). Die zunehmende digitale Durchdringung der Gesellschaft mit dieser Technologie und die damit einhergehenden Fragen sind bislang noch nicht ausreichend geklärt.

Zitierte Literatur

- Jan-Peter Ferdinand, Tobias Jetzke (2017) Voice Computing—allgegenwärtige Spracherkennung. Themenkurzprofil Nr. 15. Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag (TAB), <https://www.tab-beim-bundestag.de/de/pdf/publikationen/themenprofile/Themenkurzprofil-015.pdf>.
- Jide S. Edu, Jose M. Such, Guillermo Suarez-Tangil (2020) Smart Home Personal Assistants: A Security and Privacy Review. Preprint: Accepted for publication in ACM Computing Surveys arXiv:1903.05593v3.
- Peters, R. (2022), Sprich mit mir! Perspektiven für den Einsatz KI-basierter Dialogsysteme, Themenkurzprofil Nr. 52, Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag (TAB), <publikationen.bibliothek.kit.edu/1000143462>.
- Schaber, F., Krieger-Lamina, J., Peissl, W. (2019). Digitale Assistenten – Endbericht (p. 77). Hrsg. Institut für Technikfolgenabschätzung (ITA) der Österreichischen Akademie der Wissenschaften, Wien. doi:10.1553/ITA-pb-2019-01.

¹ advertiseonbing-blob.azureedge.net/blob/bingads/media/insight/whitepapers/2019/04%20apr/voice-report/bingads_2019voicereport.pdf.

² bloomberg.com/news/articles/2019-04-10/is-anyone-listening-to-you-on-alexa-a-global-team-reviews-audio; golem.de/news/sprachsteuerung-das-erleben-mitarbeiter-beim-abhoeren-von-siri-mitschnitten-1908-143101.html; golem.de/news/skype-und-cortana-bei-microsoft-hoeren-menschen-bei-telefonsex-mit-1908-143065.html.