Hochschule München University of Applied Sciences

Bachelorarbeit

zur Erlangung des akademischen Grades

Bachelor of Engineering (B. Eng.)

im Studiengang Technische Redaktion und Kommunikation

Analyse des Potenzials von Sprachassistenten mit Anwendungsfällen in der Technischen Kommunikation

Analysis of Potential of Voice Assistants with Use Cases in Technical Communication

---

Verfasserin:	Natascha Grundner
Matrikelnummer:	03104917
Betreuerin:	Prof. Dr. Gertrud Grünwied
Semester:	Wintersemester 2020/21
Abgabe:	30. April 2021

Zusammenfassung

Thema:

Analyse des Potenzials von Sprachassistenten mit Anwendungsfällen in der Technischen Kommunikation

Kurzfassung:

Mit der fortschreitenden Entwicklung der Künstlichen Intelligenz haben digitale Sprachassistenten stark an Bedeutung gewonnen. Der Trend geht immer mehr in Richtung sprachgesteuerte Benutzerschnittstellen (Voice User Interfaces). Ein Grund für die Popularität stellt die natürlich-sprachliche Kommunikation zwischen Mensch und Computer dar. Durch Technologien, wie Natural Language Processing, können Sprachassistenten weitaus mehr als einfache Fragen zu beantworten. Insbesondere im Bereich der Technischen Kommunikation bieten Sprachassistenten durch die intuitive und schnelle Bedienung einen Mehrwert für Anwender und kommen in verschiedenen Anwendungsfällen zum Einsatz.

In diesem Rahmen befasst sich die vorliegende Bachelorarbeit mit den Potenzialen von Sprachassistenten und Anwendungsfällen in der Technischen Kommunikation. Dabei stehen die Anwender im Fokus. Anhand von Fachliteratur wird die Nutzung von Sprachassistenten verdeutlicht und mit Studien ergänzt. Darüber hinaus wird die technische Funktionsweise mit einem anschließenden Dialogablauf eines cloudbasierten Sprachassistenten vorgestellt. Im weiteren Verlauf werden Stärken und Schwächen von Sprachassistenten näher beleuchtet. Anschließend wird eine Produkt- und Funktionsanalyse durchgeführt am Beispiel des Sprachassistenten Alexa mit der Produktvorstellung eines Smart Displays. Bei den Anwendungsfällen werden Einsatzszenarien aus den Bereichen Fremdsprachenübersetzung, Technische Dokumentation und Kundenservice aufgeführt. Zudem werden aus fünf durchgeführten Experteninterviews weitere Erkenntnisse hinsichtlich der Potenziale sowie Anwendungsfälle von Sprachassistenten in der Technischen Kommunikation generiert. Die Ergebnisse aus der Literaturarbeit und die Erkenntnisse der Experteninterviews zeigen, dass die freihändige Art der Kommunikation große Potenziale in der Technischen Kommunikation aufweist.

Abstract

Topic:

Analysis of Potential of Voice Assistants with Use Cases in Technical Communication

Abstract:

With the progressive development of artificial intelligence, digital voice assistants have become increasingly important. The trend is moving towards to voice user interfaces. One reason for this popularity is the natural language communication between users and computers. Through technologies such as Natural Language Processing, voice assistants have more functions than answering simple questions. Especially in the field of technical communication, voice assistants offer added value for users due to their intuitive and fast operation and are used in various applications.

In this context, this bachelor thesis deals with the potentials of voice assistants with use cases in technical communication. The focus is on users. Based on technical literature, the use of voice assistants is clarified and supplemented with studies. In addition, the technical functionality is presented with a subsequent dialog flow of a cloud-based voice assistant. In the further course, the strengths and weaknesses of voice assistants are examined in more detail and a product and function analysis are carried out using the example of the voice assistant Alexa with the product presentation of a specific Smart Display. The use cases include scenarios from the fields of foreign language translation, technical documentation and customer service. In addition, five interviews with experts will provide further insights into the potential and use cases of voice assistants in technical communication. The results of the literature review and the findings of the expert interviews show that hands-free communication has significant potential in technical communication.

Erklärung i. S. des § 35 Abs. 7 RaPO

Ich, Natascha Grundner erkläre, dass ich die vorliegende Bachelorarbeit selbständig verfasst, noch nicht anderweitig für Prüfungszwecke vorgelegt, keine anderen als die angegebenen Quellen oder Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate als solche gekennzeichnet habe.

Steinhöring, den 30. April 2021

Inhaltsverzeichnis

Zusammenfassung

Abstract

Erklärung i. S. des § 35 Abs. 7 RaPO

Inhaltsverzeichnis

Abkürzungsverzeichnis

Lesehinweis

1. Einleitung 1

1.1. Forschungsfragen 1

1.2. Hypothesen 2

1.3. Zielsetzung 2

1.4. Aufbau der Bachelorarbeit 3

1.5. Methoden 4

2. Technische Kommunikation 5

2.1. Mensch-Computer-Interaktion 5

2.2. Sprachsteuerung 7

2.3. User-centered Design 7

2.4. Zusammenfassung 8

3. Definitionen 9

3.1. Conversational User Interface (CUI) 9

3.2. Voice User Interface (VUI) 10

3.3. Chatbot 10

3.4. Sprachassistent 10

3.5. Zusammenfassung 12

4. Nutzung von Sprachassistenten 13

4.1. Gerätenutzung 13

4.2. Einsatzgebiete 13

4.3. Zielgruppe 15

4.4. Zusammenfassung 16

5. Technische Funktionsweise von Sprachassistenten 17

5.1. Automated Speech Recognition (ASR) 17

5.2. Natural Language Processing (NLP) 18

5.3. Technischer Dialogablauf 19

5.4. Zusammenfassung 22

6. Stärken und Schwächen von Sprachassistenten 23

6.1. Stärken von Sprachassistenten 23

6.1.1. Bewegungsfreiheit und Ortsunabhängigkeit 23

6.1.2. Intuitiv 23

6.1.3. Schnelligkeit 24

6.1.4. Sprachsuchen 24

6.1.5. Gezielte Sprachausgabe 24

6.1.6. Verfügbarkeit 25

6.1.7. Intelligenz 25

6.2. Schwächen von Sprachassistenten 26

6.2.1. Fehleranfälligkeit 26

6.2.2. Informationseinschränkung 27

6.2.3. Umgebungsgeräusche 27

6.2.4. Öffentlichkeit 27

6.2.5. Datenschutz 28

6.3. Zusammenfassung 30

7. Funktionsanalyse und Produktbeschreibung des Sprachassistenten Alexa 31

7.1. Grundlegender Funktionsumfang 31

7.2. Alexa-Skills 32

7.3. Produktbeschreibung eines Smart Displays 33

7.3.1. Technische Details 33

7.3.2. Weitere Funktionen von Smart Displays 34

7.4. Zusammenfassung 36

8. Anwendungsfälle von Sprachassistenten in der Technischen Kommunikation 37

8.1. Fremdsprachenübersetzung 37

8.2. Technische Dokumentation 38

8.3. Kundenservice 41
8.4. Zusammenfassung 43
9. Experteninterviews 44
9.1. Vorstellung der Experten 44
9.2. Vorgehen der Experteninterviews 47
9.3. Aufbau der Fragenbögen 47
9.4. Auswertung der Fragebögen 47
9.5. Zusammenfassung der ausgewerteten Experteninterviews 48
9.5.1. Zielgruppe von Sprachassistenten 48
9.5.2. Stärken und Nutzung von Sprachassistenten 48
9.5.3. Schwächen von Sprachassistenten 49
9.5.4. Anwendungsmöglichkeiten in der Technischen Kommunikation 50
9.5.5. Zukunftsblick 52
9.6. Zusammenfassung 53
10. Fazit und Ausblick 54
10.1. Fazit 54
10.2. Ausblick 56

Abbildungsverzeichnis

Tabellenverzeichnis

Literaturverzeichnis

Anhang

Abkürzungsverzeichnis

Abkürzung	Begriff
API	**A**pplication **P**rogramming **I**nterface
ASR	**A**utomatic **S**peech **R**ecognition
DIN	**D**eutsches **I**nstitut für **N**ormung
ebd.	**eb**en**d**a
et al.	**Et al**ii
GUI	**G**raphical **U**ser **I**nterface
HCI	**H**uman **C**omputer **I**nteraction
IoT	**I**nternet **o**f **T**hings
ISO	**I**nternationale **O**rganisation für Normung
KI	**K**ünstliche **I**ntelligenz
NLG	**N**atural **L**anguage **G**eneration
NLP	**N**atural **L**anguage **P**rocessing
NLU	**N**atural **L**anguage **U**nderstanding
RGB	**R**ot, **G**rün und **B**lau
STT	**S**peech-**T**o-**T**ext
TTS	**T**ext-**T**o-**S**peech
VUI	**V**oice **U**ser **I**nterface

Lesehinweis

Aus Gründen der besseren Lesbarkeit wird in dieser Bachelorarbeit die Sprachform des generischen Maskulinums angewandt. Sämtliche Ausführungen sind jedoch geschlechtsunabhängig und beziehen sich damit auf alle Geschlechter.

1. Einleitung

„Der Einsatz von neuen Technologien ist vergleichbar mit dem Einsatz von Werkzeugen. Richtig eingesetzt sollen sie unterstützen, helfen und das Leben erleichtern.“[1]

Die intuitiv bedienbaren Assistenzsysteme haben sich als fester Bestandteil im Alltag etabliert und bieten zahlreiche Anwendungsmöglichkeiten. Ein Sprachassistent unterstützt die barrierefreie Nutzung und ermöglicht vielen Personen Zugänglichkeit zur Technik. Die Künstliche Intelligenz ist in der Lage, die Intention des Anwenders zu verstehen und mit Hilfe von Wissensdatenbanken eine auf den Anwender passgenaue Antwort zu finden.[2]

Sprachassistenten erschließen neue Anwendungsfelder. Insbesondere im Bereich der Technischen Kommunikation kann die sprachliche Schnittstelle einen Mehrwert für Anwender bieten und in vielen Anwendungsfällen zum Einsatz kommen. Die Vorstellung, dass Sprachassistenten autonom Anrufe entgegennehmen, Termine vereinbaren und die Anrufer dabei nicht erkennen, ob es sich um eine menschliche oder synthetische Stimme handelt, zeigt, dass Sprachassistenten in Zukunft große Potenziale bieten.

Doch wie beim Aufkommen jeder neuen Technologie steht die Künstliche Intelligenz vor Herausforderungen. Stellt die Spracheingabe eine sinnvolle Interaktionsmöglichkeit zwischen der Mensch-Computer-Schnittstelle dar? Welche Stärken und Schwächen sind damit verbunden? Inwieweit unterstützen intuitive Spracheingaben bei der Verwendung von Bedienungsanleitungen? Können Sprachassistenten Mitarbeiter im Kundenservice ersetzen? Diese und weitere Fragen werden im Laufe der Bachelorarbeit beantwortet.

1.1. Forschungsfragen

Folglich befasst sich die Bachelorthesis „Analyse der Potenziale von Sprachassistenten mit Anwendungsfällen in der Technischen Kommunikation“ mit diesen Forschungsfragen:

F1: Wie ist die technische Funktionsweise von Sprachassistenten aufgebaut?

F2: Bieten Sprachassistenten in der Technischen Dokumentation Vorteile?

[1] M. Kaiser et al. (2019) S. 11

[2] Vgl. T. Hörner (2019) S. 7 f.

F3: Inwieweit können Sprachassistenten im Kundenservice eingesetzt werden?

F4: Welche Chancen bieten Sprachassistenten in der Fremdsprachenübersetzung?

1.2. Hypothesen

Anhand der aufgeführten Forschungsfragen ergeben sich folgende Hypothesen, die im Rahmen der Bachelorarbeit belegt bzw. widerlegt werden:

Hypothese	Forschungsfrage
H1: Natural Language Processing gilt als Schlüsseltechnologie für Sprachassistenten.	**F1**
H2: Anwender gelangen mit intuitiven Spracheingaben in der Technischen Dokumentation schnell und gezielt an Informationen.	**F2**
H3: Bei der Vermittlung von Handlungsanweisungen eignen sich Smart Displays.	**F2**
H4: Mit Sprachassistenten können Mitarbeiter im Kundenservice ersetzt werden.	**F3**
H5: Mit Sprachassistenten ergeben sich neue Möglichkeiten in der Fremdsprachenübersetzung.	**F4**

Tabelle 1: Hypothesen und zugehörige Forschungsfragen[3]

1.3. Zielsetzung

Sprachassistenten können durch die intuitive Bedienung und intelligente Auswertung von Daten einen Mehrwert in der Technischen Kommunikation darstellen. Ziel der Bachelorarbeit ist es, diese Potenziale zu beleuchten und Anwendungsfälle in der Technischen Kommunikation darzustellen. Die Bachelorarbeit baut auf einer Literaturrecherche auf. Zudem erfolgt eine Produkt- und Funktionsanalyse am Beispiel des

[3] Eigene Darstellung

Sprachassistenten Alexa. Als qualitative Forschungsmethode werden fünf Experteninterviews durchgeführt. Die Experten werden entsprechend ihres Fachbereiches zu Potenzialen und Anwendungen von Sprachassistenten in der Technischen Kommunikation befragt. Auf Basis der Ergebnisse aus der Literaturarbeit und den Erkenntnissen der Experten werden im Fazit Rückschlüsse gezogen, inwieweit die aufgestellten Hypothesen belegt werden können.

1.4. Aufbau der Bachelorarbeit

Auf Basis der Zielsetzung setzt sich die Bachelorarbeit aus folgenden Hauptbestandteilen zusammen:

- Literaturarbeit
- Produkt und Funktionsanalyse des Sprachassistenten Alexa
- Experteninterviews

Zu Beginn leitet das Kapitel 2 die Entwicklung der Mensch-Computer-Interaktion ein. Dabei wird der Stellenwert der Sprachsteuerung in der Technischen Kommunikation erläutert. Das Kapitel 3 beschäftigt sich mit Begrifflichkeiten aus dem Bereich Conversational-User-Interfaces. Im Anschluss gibt das Kapitel 4 Einblick über die Nutzung und Einsatzgebiete von Sprachassistenten und beschreibt anschließend die Zielgruppe.

Für das Verständnis der technischen Funktionsweise befasst sich das Kapitel 5 zum einen mit den Kerntechnologien der Sprachverarbeitung und zum anderen mit einem Dialogablauf. Jede Technologie weist Chancen als auch Risiken auf. Um diese zu beleuchten, behandelt das Kapitel 6 die Stärken und Schwächen von Sprachassistenten.

Das Kapitel 7 spezialisiert sich auf einen Sprachassistenten und stellt eine Produkt- und Funktionsanalyse dar. Für die Analyse wird sich auf den cloudbasierten Sprachassistenten Alexa von Amazon festgelegt. Nach einem Überblick des Funktionsumfangs, folgt eine Produktbeschreibung eines Smart Displays. Dieser bietet weitere nützliche Funktionen für Anwender.

Das Kapitel 8 beschreibt Anwendungsfälle von Sprachassistenten in der Technischen Kommunikation. In diesem Kapitel wird erläutert, ob Sprachassistenten einen Mehrwert für Anwender bieten in den Bereichen Fremdsprachenübersetzung, Technischen Dokumentation und Kundenservice. Das Kapitel 9 befasst sich mit dem empirischen Teil der Bachelorarbeit. Abschließend rundet das Kapitel 10 die Bachelorarbeit mit einem Fazit und einem Ausblick ab und geht auf die vorangestellten Hypothesen ein, die verifiziert bzw. falsifiziert werden.

1.5. Methoden

Die Analyseergebnisse setzen sich aus folgenden Methoden zusammen:

Literaturarbeit: Zu Beginn der vorliegenden Bachelorarbeit werden theoretische Grundlagen und der aktuelle Forschungsstand vermittelt. Hierzu wird als Methode eine Literaturarbeit angewandt. Dabei wird sich auf Fachliteratur bezüglich Conversational User Interfaces und Sprachassistenten sowie auf verschiedenen Studien berufen.

Produkt- und Funktionsanalyse am Beispiel des Sprachassistenten Alexa: Nachdem der grundlegende Teil in der Literaturarbeit erläutert wurde, wird in der zweiten Methode eine Produkt-und Funktionsanalyse des Sprachassistenten Alexa durchgeführt. Diese Analyse gilt als Bestandsaufnahme des Sprachassistenten und gewährt einerseits einen Überblick über den Funktionsumfang mit der Erweiterung von Voice-Apps und andererseits wird ein Smart Display beschrieben.

Experteninterviews: Im empirischen Teil werden fünf Experteninterviews in schriftlicher Form durchgeführt. Dabei werden Experten aus unterschiedlichen Fachbereichen zu den Potenzialen und Anwendungsfällen von Sprachassistenten befragt.

2. Technische Kommunikation

Die Digitalisierung und Internet of Things (IoT) tragen zu einem Wandel der Technischen Kommunikation bei.

„Technische Kommunikation ist der Prozess der Definition, Erstellung und Bereitstellung von Informationsprodukten für die sichere, effiziente und effektive Verwendung von Produkten (technische Systeme, Software, Dienstleistungen).“[4]

Intelligente Technologien, wie Smart City, Smart Factory oder Smart Home rücken in den Vordergrund. Zunehmend mehr Geräte werden vernetzt und können eine hohe Datenmenge in Echtzeit verarbeiten.[5] Mit der Weiterentwicklung neuer Technologien verändern sich die Interaktionsmöglichkeiten mit technischen Geräten. Anfangs waren Maus und Tastatur eine gängige Eingabemöglichkeit. Jetzt rückt die Sprachsteuerung immer mehr in den Fokus als natürliche Interaktion des Menschen mit der Technik.

2.1. Mensch-Computer-Interaktion

Die Mensch-Computer-Interaktion beschäftigt sich mit der Frage, wie der Umgang von Anwendern mit softwarebasierten Anwendungen erfolgt. Im Mittelpunkt stehen neue Technologien, die effizient eingesetzt werden sollen.[6]

Herr Hellige betont in seinem Werk „Mensch-Computer-Interface“, dass neben technischen Funktionalitäten bei Mensch-Computer-Schnittstellen, verstärkt die psychologischen, ergonomischen und sozialen Kontexte der Anwender von hoher Bedeutung sind. Das Interface veranschaulicht der Autor als eine Brücke zwischen Anwender und dem technischen System. Diese Schnittstelle ist geprägt von Vorerfahrungen der Anwender und Entwickler. Mit Hinblick auf die Entwicklung der Mensch-Computer-Interaktion, mussten sich Anwender am Anfang an den Computer anpassen. Heutzutage stehen die menschlichen Bedürfnisse im Mittelpunkt.[7]

Die folgende Abbildung stellt die Entwicklung der Mensch-Computer-Schnittstellen im Laufe der Zeit dar.

---

[4] Tekom: https://www.tekom.de/technische-kommunikation-das-fach/was-ist-technische-kommunikation

[5] Vgl. Tekom: https://www.tekom.de/technische-kommunikation-das-fach/intelligente-information

[6] Vgl. M. Flückiger, M. Richter (2016) S. 7 f.

[7] Vgl. H. Hellige (2008) S. 7

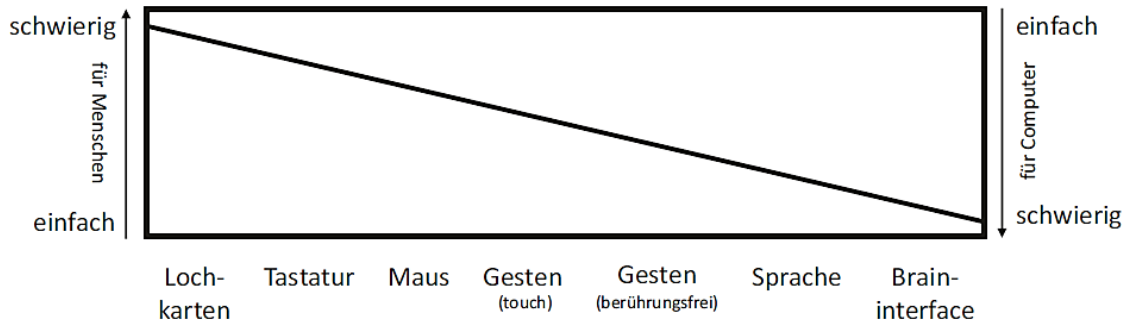

Abbildung 1: Entwicklung der Mensch-Computer-Schnittstellen[8]

Die Entwicklung zeigt, dass mit jeder neuen zunehmend komplexeren Technologie, die Bedienung für Anwender einfacher wird. Die Computernutzung hat in den 1950er-Jahren mit Lochkarten angefangen. Papierstreifen oder einzelne Karten aus Karton sind mit einem Lochcode versehen worden. Ausschließlich Spezialisten waren fähig diesen Code zu interpretieren. Eine erhebliche Vereinfachung der Computernutzung, brachte die Einführung der grafischen Benutzerschnittstelle mit Tastatur und Maus mit sich. Dadurch konnten viele Personen Computer bedienen.

Danach folgte der Touchscreen mit Gestensteuerung. Eine erstmals intuitive und ortsunabhängige Bedienung, die vor allem mit der Einführung des ersten iPhones von Apple im Jahr 2007 für Aufsehen sorgte. Die Sprachsteuerung gilt als neuer technologischer Fortschritt, die in immer mehr technischen Geräten integriert wird.

Die neueste Technologie ist das sogenannte „Brain Interface“, das eine Kommunikation zwischen einem menschlichen Gehirn und einem technischen System ermöglicht. Als neurotechnologisches Eingabesystem können mentale Befehle sprach- und bewegungsunabhängig Systeme steuern. Diese Art von Interaktionsform steht aber derzeit noch vor einer technologischen Herausforderung und muss noch weiter erforscht werden.[9]

[8] T. Hörner (2019) S. 2

[9] Vgl. C. Heuer: https://www.int.fraunhofer.de/de/geschaeftsfelder/corporate-technology-foresight/Brain-Computer-Interfaces.html

2.2. Sprachsteuerung

„Sprache ist die natürlichste Kommunikationsform.“[10]

Dadurch ergibt sich ein hohes Potenzial für Sprachassistenten, bei denen die Interaktion durch natürliche Sprache erfolgt. Je mehr Geräte mit Sprachassistenten ausgestattet sind, desto natürlicher ist der Umgang für Anwender mit Geräten zu sprechen.[11]

Heutzutage sind Sprachassistenten omnipräsent und unterstützen in unterschiedlichen Lebensbereichen. Sie können weitaus mehr als Musik abzuspielen, Nachrichten abzurufen oder Timer zu stellen. Die Künstliche Intelligenz bietet für Anwender großes Potenzial mit vielen unterstützenden Funktionen.

Für die Technische Kommunikation hat die Sprachsteuerung eine hohe Relevanz. Mit den Anwendungsfällen wird sich im weiteren Verlauf der Bachelorarbeit beschäftigt.

2.3. User-Centered Design

Wie das Kapitel 2.1 zeigt, sind die Bedürfnisse der Anwender in der Mensch-System-Interaktion von hoher Wichtigkeit. Um diese zu erfüllen, ist die benutzerorientierte Gestaltung bedeutsam.

User-Centered Design ist ein wichtiger Bestandteil der Softwareentwicklung und hängt eng mit Usability, User-Experience sowie Barrierefreiheit zusammen. Dabei steht die benutzerorientierte Gestaltung der Software mit hoher Nutzerfreundlichkeit im Vordergrund.[12] Als Standard der Mensch-System-Interaktion gilt die europäische Normenreihe EN ISO 9241 „Ergonomie der Mensch-System-Interaktion“. Darunter werden Anforderungen für das Design ergonomischer Benutzerschnittstellen bestimmt.

Der Teil 210 der Normenreihe trägt den Titel „Menschzentrierte Gestaltung interaktiver Systeme. In dieser Norm werden Bestimmungen festgelegt sowie Grundsätze zur Gestaltung für den Lebenszyklus interaktiver Systeme empfohlen, wie z.B. Software- und Hardwarebestandteile die Mensch-System-Interaktion verbessern könnten.[13]

---

[10] P. Kabel (2020) S. 1

[11] Vgl. Bundesverband Digitale Wirtschaft (BVDW) e.V. (April 2019): https://www.bvdw.org/fileadmin/bvdw/upload/publikationen/smart_home/Sprachassistenten_im_SmartHome.pdf

[12] Vgl. Ryte Wiki: https://de.ryte.com/wiki/User_Centered_Design

[13] Vgl. DIN EN ISO 9241-210:2020-03: https://www.beuth.de/de/norm/din-en-iso-9241-210/313017070

2.4. Zusammenfassung

Wie in den vorherigen Kapiteln beschrieben, hat sich durch die Weiterentwicklung neuer Technologien die Möglichkeiten der Interaktionen mit technischen Geräten verändert. Heutzutage spielt die Sprachsteuerung eine elementare Rolle zwischen Mensch und Computer. Besonders Sprachassistenten, die auf Künstlicher Intelligenz basieren, bieten im Bereich der Technischen Kommunikation hohes Potenzial mit vielen Anwendungen. Die Voraussetzung liegt hierbei, dass Sprachassistenten stets auf die Bedürfnisse der Anwender ausgerichtet sind.

3. Definitionen

Dieses Kapitel befasst sich mit Begrifflichkeiten rund um Sprachassistenten.

3.1. Conversational User Interface (CUI)

Die Begriffsdefinition für ein CUI umfasst die Kommunikation zwischen Mensch und Computer in gesprochener sowie in geschriebener Sprache. Beispiele für CUIs sind Chatbots und Sprachassistenten.[14] Dabei ermöglichen Schnittstellen, wie z.B. Application Programming Interfaces (APIs), den Zugriff auf Informationen aus verschiedenen Wissensdatenbanken.[15] Diese Möglichkeit der intelligenten Interaktion verbindet Daten, Anwendungen, Prozesse und Endgeräte und ist Basis einer vernetzten Welt.[16]

Der Begriff „Multimodalität“ beschreibt die Integration von CUIs in unterschiedlicher Hardware. In vielen Situationen ist es besser, Informationen zusätzlich zu sehen, statt nur den auditiven Wahrnehmungskanal anzusprechen. Aus diesem Grund fungieren viele Geräte besser mit Kombination eines Bildschirms, wie zum Beispiel Smart Displays. Diese Produktart wird näher im Kapitel 7.3. beleuchtet.

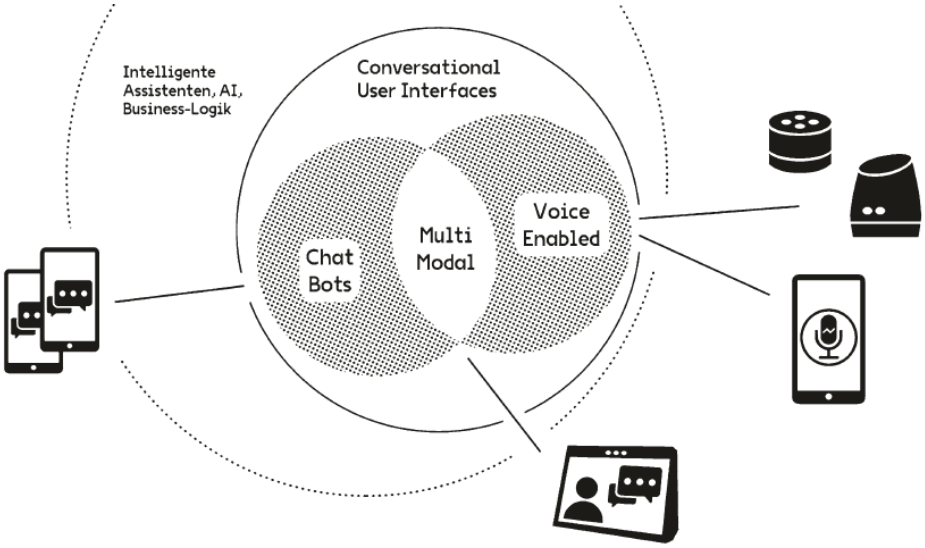

Abbildung 2: Integrierte CUIs in unterschiedlicher Hardware[17]

[14] Vgl. P. Kabel (2020) S. 1

[15] Vgl. P. Kabel (2020) S. 145

[16] Vgl. B. Leukert (2019) S. 46

[17] P. Kabel (2020) S. 3

3.2. Voice User Interface (VUI)

Bei einem Voice User Interface erfolgt die Interkation zwischen Mensch und Computer über Spracheingaben und –ausgaben mit Hilfe von automatischer Spracherkennung und -synthese.[18] Bei der Konzeption sollte die Nutzerperspektive eingenommen werden, um ein positives Nutzererlebnis zu ermöglichen.

VUIs bringen deutliche Vorteile mit sich. Die Anwender benötigen bei der Bedienung eines VUIs ausschließlich ihre Stimme und müssen nicht tippen, scrollen oder klicken. Bei der Interaktion mit einer sprachbasierten Benutzerschnittstelle steht der auditive Wahrnehmungskanal im Vordergrund.[19]

3.3. Chatbot

Unter Chatbots werden technische Dialogsysteme bezeichnet mit einer text- oder sprachbasierten Kommunikation. Sie treten als animierter oder statischer Avatar auf Unternehmenswebseiten oder in Instant-Messaging-Diensten auf.[20]

Chatbots, die anfangs rein textbasiert arbeiteten, haben sich zu intelligenten Assistenten entwickelt, die mit natürlicher Sprache kommunizieren. Sie stellen eine hohe Wichtigkeit für den Kundenservice dar. Unternehmen setzen Chatbots als Kommunikationspartner für Kunden ein. Beispielsweise für die Produktberatung, für Hilfestellungen bei Problemen oder für die Ausführung von einfachen Kundenanliegen.

Die Funktionsweise eines Chatbots basiert auf Künstlicher Intelligenz und maschinellem Lernen. Daher ist er in der Lage mit Hilfe von Training und Feedbackanalyse Kundenanliegen zu verstehen und diese auszuführen. Heutzutage ist es möglich, dass Chatbots sprachliche Antworten ausgeben, die Anwender oft nicht von Servicemitarbeitern unterscheiden können.[21]

3.4. Sprachassistent

Ein Sprachassistent ist eine Software im Rechenzentrum von Anbietern und kann auf unterschiedlichen Geräten verwendet werden. Um auf das Rechenzentrum zuzugreifen, wird eine aktive Internetverbindung benötigt. Um die Spracheingaben der Anwen-

[18] Vgl. P. Kabel (2020) S. 201

[19] Vgl. Ryte Wiki: https://de.ryte.com/wiki/User_Interface#Voice_User_Interface_VUI

[20] Vgl. O. Bendel (19. Februar 2018): https://wirtschaftslexikon.gabler.de/definition/chatbot-54248

[21] Vgl. S. Luber, N. Litzel (Februar 2018): https://www.bigdata-insider.de/was-ist-ein-chatbot-a-690591/

der zu bearbeiten, beziehen sich Sprachassistenten auf Big Data und Formen der Künstlichen Intelligenz. Sie sind in mobilen Geräten als integrierte Funktion oder als App verfügbar sowie auf stationären Geräten, wie z.B. Smart Speaker oder Smart Displays.[22]

In dem von „Voicebot“ im Jahr 2018 veröffentlichten Bericht „Voice Assistant Consumer Adoption“ wird die Entwicklung von Sprachassistenten in zwei Phasen dargestellt.

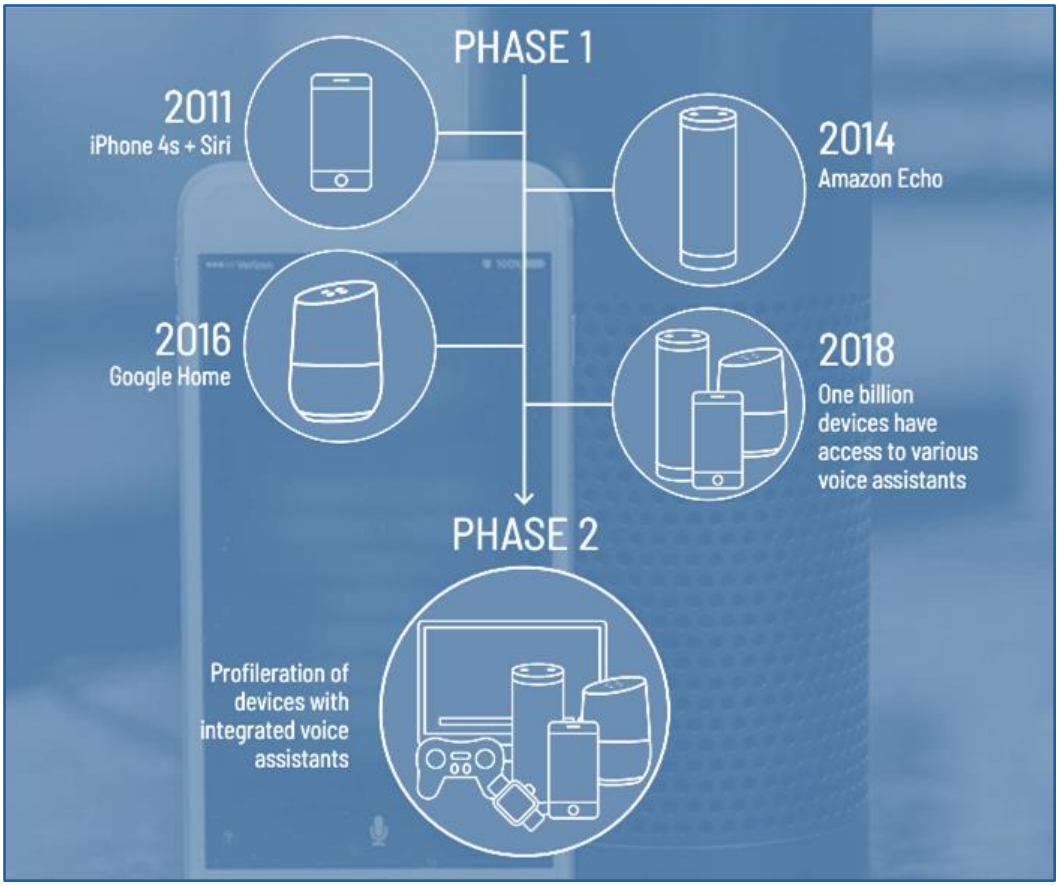

Abbildung 3: Geräteentwicklung mit integrierten Sprachassistenten[23]

In der ersten Phase gab es vor allem erste Berührungspunkte mit der Sprachsteuerung am Smartphone mit der Einführung des implementierten Sprachassistenten Siri im iPhone 4s von Apple im Jahr 2011. Danach folgte im Jahr 2014 der Sprachassistent Alexa im stationären Lautsprecher „Amazon Echo“ und im Jahr 2016 der Sprachassistent Google Assistant im Lautsprecher „Google Home“. Vier Jahre später hatten rund eine Milliarde Geräte Zugang zu verschiedenen Sprachassistenten. In der Phase zwei

[22] Vgl. Ryte Wiki: https://de.ryte.com/wiki/Digitale_Sprachassistenten

[23] Vgl. B. Kinsella, A. Mutchler (November 2018): https://voicebot.ai/wp-content/uploads/2019/01/voice-assistant-consumer-adoption-report-2018-voicebot.pdf

können Anwender mit der Sprachsteuerung eine Vielzahl von unterschiedlichen Geräten bedienen, wie zum Beispiel Smart Displays, Smartwatches oder Smart-TVs.[24]

3.5. Zusammenfassung

CUIs, VUIs, Chatbots und Sprachassistenten spielen in der heutigen Zeit eine elementare Rolle. Sprachassistenten können in unterschiedlicher Hardware integriert werden und sind vielfältig einsetzbar. Zugleich rückt der Begriff „Multimodalität“ immer näher. Das Zusammenspiel zwischen Sprachassistent und Bildschirm bietet Potenzial, denn Anwender erhalten durch die Ergänzung eines Bildschirmes zusätzliche Informationen.

[24] Vgl. B. Kinsella, A. Mutchler (November 2018): https://voicebot.ai/wp-content/uploads/2019/01/voice-assistant-consumer-adoption-report-2018-voicebot.pdf

4. Nutzung von Sprachassistenten

Dieses Kapitel gilt der näheren Betrachtung der generellen Nutzung von Sprachassistenten. Zuerst wird untersucht über welche Geräte und für welche Anwendungen Sprachassistenten häufig genutzt werden. Danach wird sich mit der Zielgruppe von Sprachassistenten beschäftigt.

4.1. Gerätenutzung

Die Sprachsteuerung führt zu einer zunehmenden Beliebtheit. Dies bestätigt die Umfrage des Digitalverbands Bitkom, die im Mai 2020 durchgeführt wurde. Zwei von fünf Internetnutzern in Deutschland verwenden geräteübergreifend die Sprachsteuerung. Unter den 16- bis 29-jährigen Internetnutzern sind es 59 Prozent.

Wie in dem Balkendiagramm dargestellt ist, verwenden zwei Drittel der Anwender Sprachassistenten am Smartphone. Auf dem zweiten Platz liegt das Tablet mit 41 Prozent, gefolgt von Smart Speaker mit 26 Prozent und Smart-TVs mit 21 Prozent. Am wenigsten Verwendung finden derzeit noch Smartwatches, Kopfhörer und Smart Displays die jeweils unter 10 Prozent liegen.[25]

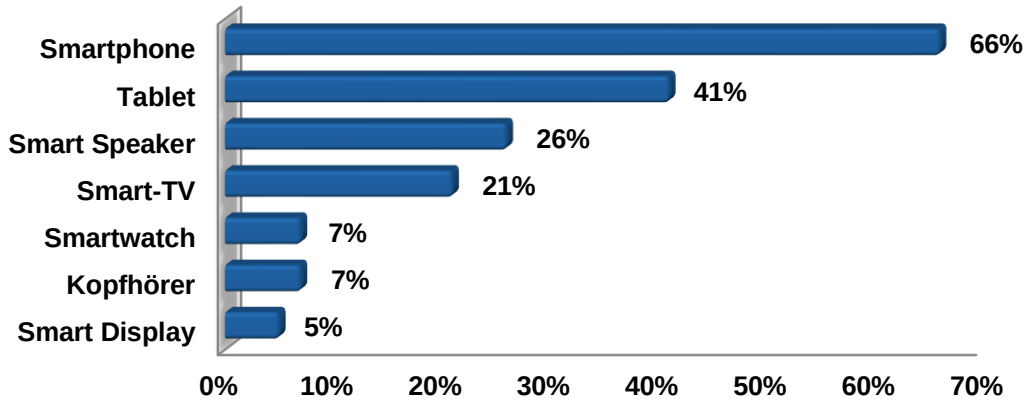

Abbildung 4: Nutzung von Sprachassistenten[26]

4.2. Einsatzgebiete

Die hohe Reichweite von Sprachassistenten ermöglicht ein facettenreiches Feld an Anwendungsmöglichkeiten im privaten als auch im wirtschaftlichen Bereich. Monitor

[25] Vgl. C. Krösmann, S. Klöß (24. Juli 2020): Beliebte Helfer: Sprachassistenten haben sich durchgesetzt

[26] Vgl. ebd.

Deloitte befasste sich mit dem Thema, wie in der Zukunft sprachbasierte Benutzeroberflächen den Handel in Europa beeinflussen werden und veröffentlichte eine Studie namens „Beyond Touch: Voice-Commerce 2030“.[27] Die nachfolgende Abbildung zeigt, wie vielfältig sich das Anwendungsfeld von Sprachassistenten gestaltet.

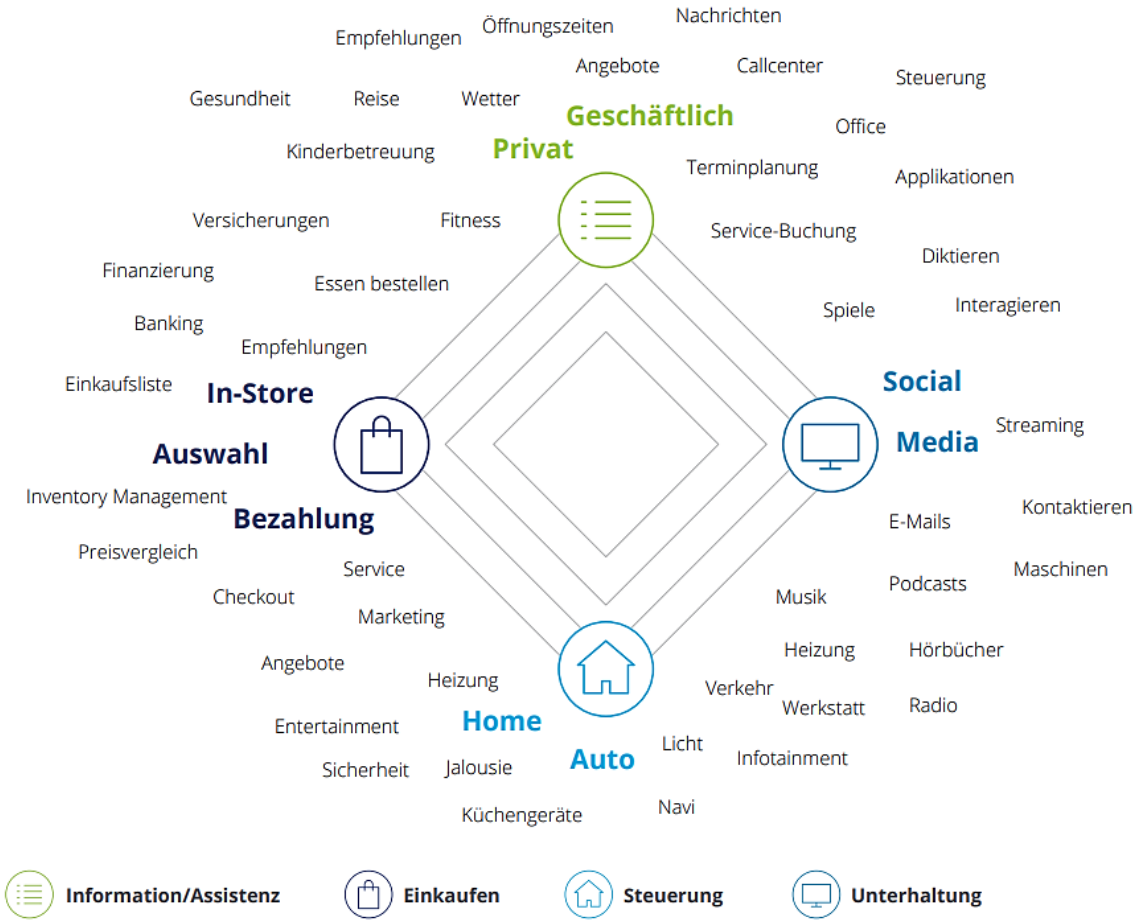

Abbildung 5: Einsatzgebiete von Sprachassistenten[28]

Wie in der Abbildung dargestellt ist, klassifizieren sich die vier wichtigsten Einsatzgebiete von Sprachassistenten in:

- Information bzw. Assistenz (Terminplanung, Nachrichten etc.)
- Einkaufen (Service, Preisvergleich etc.)
- Steuerung (Smart Home und Fahrzeug)
- Unterhaltung (Musik, Spiele, Hörbücher etc.)

[27] Vgl. Monitor Deloitte: Beyond Touch – Voice Commerce 2030 (November 2018): https://www2.deloitte.com/de/de/pages/consumer-business/articles/sprachassistent.html

[28] Vgl. ebd.

Laut der Studie werden Sprachassistenten im privaten Gebrauch häufig dafür genutzt, um Informationen abzurufen, wie sich über die aktuellen Nachrichten oder das Wetter zu informieren. Zudem werden Sprachassistenten häufig im Bereich der Unterhaltung genutzt. Dazu zählt die Steuerung von Musik oder das Abspielen von Filmen und Hörbücher. Daneben ist ein weiterer beliebter Einsatz von Sprachassistenten im Smart Home. Per Sprache werden Heizungen, Lichter oder Jalousien bedient. Darüber hinaus stellt die Sprachsteuerung eine hohe Relevanz im Fahrzeug dar. Auf dieses Einsatzgebiet wird näher im Kapitel 8.2. eingegangen.

Die Anwendungsmöglichkeiten bewegen sich, neben den Privatgebrauch, verstärkt im geschäftlichen Bereich, beispielsweise im Kundenservice. Dieser Bereich bietet ein hohes Potenzial für Kunden und wird ausführlich im Kapitel 8.3. beschrieben. Aus der Studie geht darüber hinaus hervor, dass Voice Commerce in der Zukunft eine hohe Relevanz für den Handel spielen wird. Beispielsweise wird immer mehr sprachgesteuertes Einkaufen in den Vordergrund rücken. Sprachassistenten bieten beispielsweise Unterstützung beim Online-Shopping durch personalisierte Empfehlungen auf Basis vorheriger Käufe oder beantworten Produktfragen. Unternehmen sollten demnach Strategien entwickeln, wie deren Produkte und Services mit sprachbasierten Benutzeroberflächen integrierbar sind.[29]

4.3. Zielgruppe

Für Sprachassistenten kommen sämtliche Altersgruppen als Zielgruppe in Frage. Sowohl „Digital Natives“, die mit dem Internet aufgewachsen und mit digitalen Medien vertraut sind als auch „Digital Immigrants“, die erst im Laufe der Zeit mit Smartphones und dem Internet in Berührung gekommen sind.[30]

An einer repräsentativen Digitalstudie der Postbank im Jahr 2020 nahmen insgesamt 3.035 Teilnehmer aus Deutschland teil. Die Ergebnisse zeigen, dass 59 Prozent der Anwender unter 40 Jahren Sprachassistenten verwenden. Im Vergleich zum Jahr 2019 waren es 48 Prozent. Die Studie hat zudem herausgefunden, dass tendenziell Familien Alexa & Co. nutzen als alleinlebende Menschen. Im Vergleich zu 59 Prozent der Haushalte ab vier Personen, nutzen nur ein Drittel der Single-Haushalte aktiv Sprachassistenten.

[29] Vgl. Monitor Deloitte: Beyond Touch – Voice Commerce 2030 (November 2018): https://www2.deloitte.com/de/de/pages/consumer-business/articles/sprachassistent.html

[30] Vgl. M. Siepermann (19. Februar 2018): https://wirtschaftslexikon.gabler.de/definition/digital-native-54496

Der Leiter der digitalen Transformation der Postbank Herr Brosch ist der Meinung, dass sich in Zukunft bei Menschen in den Altersgruppen 50 und 60 plus Sprachassistenten fest etablieren werden. Grund hierfür ist, dass ein intuitiver Sprachbefehl einfacher auszuführen ist als im Menü zu scrollen und zu tippen.[31]

Sprachassistenten spielen zudem für die Barrierefreiheit eine elementare Rolle. Über die Sprachsteuerung kann der Alltag von Personen mit Sehbehinderungen oder motorischen Einschränkungen effektiv unterstützt werden. Dies gilt auch für Personen mit einer Lese- und Rechtschreibschwäche. Der Geschäftsführer des Deutschen Blinden- und Sehbehindertenverbands (DBSV) Herr Bethke erkennt bei der Technologie der Spracherkennung ein großes Potenzial für Menschen mit Seheinschränkungen.[32]

Intuitive Sprachbefehle erleichtern den Alltag bei einer Vielzahl von Aufgaben. Mit einem sprachgesteuerten Smart-Home-System lässt sich beispielsweise auf Zuruf die Beleuchtung, die Heizung oder die Klingelanlage steuern. Ein smarter Lautsprecher kann eine Einkaufsliste erstellen, einen Wecker stellen oder an Terminen erinnern, wie die Einnahme von Medikamenten.[33]

Über die sprachbasierte Benutzeroberfläche werden allerdings auch Nutzergruppen ausgeschlossen. Für Menschen mit Hör- oder Sprechbeeinträchtigungen stellt eine Sprachsteuerung eine Barriere dar.

4.4. Zusammenfassung

Sprachassistenten werden am häufigsten über Smartphones, Tablets und Smart Speaker benutzt. Das Besondere an Sprachassistenten ist die breit gefächerte Zielgruppe. Sowohl Kinder als auch ältere Menschen, die mit der Technik nicht vertraut sind, können einfach und intuitiv Sprachbefehle erteilen. Außerdem unterstützt die Sprachsteuerung die barrierefreie Nutzung. Menschen, die durch Krankheit, Alter oder Behinderungen beeinträchtigt sind, profitieren von Sprachassistenten. Die Einsatzgebiete gestalten sich als facettenreich und dringen in immer mehr Bereichen.

[31] Vgl. H. Schlegel (29. Juni 2020): https://www.presseportal.de/pm/6586/4637002

[32] Vgl. T. Heuzeroth (21. Februar 2021): https://www.welt.de/wirtschaft/webwelt/article226788181/Mit-Spracherkennung-Amazon-laesst-sich-von-Blinden-helfen.html

[33] Vgl. Pflege digital: https://www.pflege.digital/

5. Technische Funktionsweise von Sprachassistenten

Um sich ein Grundverständnis der technischen Funktionsweise von Sprachassistenten zu verschaffen, beschäftigt sich dieses Kapitel mit den Kerntechnologien, die eine funktionierende sprachbasierte Interaktion zwischen Anwender und Sprachassistenten ermöglichen. Im Anschluss wird ein Dialogablauf am Beispiel des cloudbasierten Sprachassistenten Alexa dargestellt.

5.1. Automated Speech Recognition (ASR)

Die Technologie ASR ist neben Natural Language Processing (NLP) eine wichtige Komponente für die Spracherkennung. ASR ist verantwortlich für die Erkennung von Sprache und die Umwandlung in Textform. Die Spracherkennung versteht das, wofür das System mit Algorithmen trainiert wird.[34] Durch Big Data und Deep Learning kann die Spracherkennung eine hohe Datenmenge auswerten.[35]

Die grundlegende Funktionsweise einer ASR-Software lässt sich wie folgt beschreiben:

Grundlegende Funktionsweise einer ASR-Software:
1. Umwandlung der Spracheingabe in eine WAV-Datei
2. Zerlegung der WAV-Datei in Phoneme
3. Umwandlung der Phoneme in Wörter
4. Anwendung eines Sprachmodells zur Wortkorrektur
5. Sprachausgabe

Tabelle 2: Funktionsweise der Technologie ASR[36]

Die Spracheingabe wird in eine WAV-Datei umgewandelt – ein Wellenform-Format. Dabei handelt es sich um Rohdateien, die unterschiedliche Bitraten speichern können.[37] Anschließend wird die WAV-Datei durch eine maschinell lernbasierte Analyse in

[34] Vgl. J. Schleier, H. Hauser (2020) S. 208

[35] Vgl. P. Kabel (2020) S. 190

[36] Vgl. ebd. S. 53

[37] Vgl. C. Schmidt (27. Dezember 2019): https://www.canto.com/de/blog/wav-audioformat/

sogenannte Phoneme zerlegt.[38] Ein Phonem bezeichnet die kleinste bedeutungsunterscheidende sprachliche Einheit, wie z.B. das aus drei Buchstaben bestehende „sch“.[39]

Unternehmen, wie z.B. Google, haben Audiodateien gesammelt und in Phoneme umgewandelt. Anhand dieser Lerndaten können Wörter und Sätze für neue Audiodateien übersetzt werden. Nach der Umwandlung in Phoneme, wird das wahrscheinlichste Wort erkannt, dies ein Phonem bilden kann. Beispielsweise könnte ein ASR-System folgendes ausgeben:

- HHHEE_LL_LLLOOO
- HHHUU_LL_LLLOOO
- AAAUU_LL_LLLOOO

Wiederholte Buchstaben und Leerzeichen werden bereinigt und Wörter, wie „Hello“, „Hullo“ und „Aullo“ resultieren. Das wahrscheinlichste Wort, das im Wörterbuch erkannt wird, ist „Hello“. Nachdem einzelne Wörter analysiert werden, wiederholt sich dieser Prozess mit ganzen Sätzen. Ein Sprachmodell wendet eine Wahrscheinlichkeitsanalyse an mit Basis auf gesammelten Daten, bei der festgestellt wird, wie wahrscheinlich es ist, dass ein Wort einem anderen Wort folgt.[40]

5.2. Natural Language Processing (NLP)

Die Technologie NLP gilt als Vorstufe der Künstlichen Intelligenz. Der Computer beruft sich auf künstliche erzeugte Erfahrungen und Algorithmen, die auf statistischem maschinellem Lernen basieren.[41] Je mehr der lernfähige Sprachassistent mit Sprachaufzeichnungen unterschiedlicher Anwender trainiert wird, desto besser kann sich der Sprachassistenten an Dialekten, Sprachgewohnheiten, Vokabular oder Hintergrundgeräuschen einstellen.[42]

Die Herausforderung der Technologie NLP liegt im Verständnis der menschlichen Sprache, denn neben einzelnen Wörtern, muss der Computer ganze Sätze sowie Zusammenhänge erkennen und verstehen. Des Weiteren kommt die Schwierigkeit hinzu,

[38] Vgl. P. Kabel (2020) S. 52

[39] Vgl. Duden Rechtschreibung (2021): https://www.duden.de/rechtschreibung/Phonem

[40] Vgl. P. Kabel (2020) S. 52

[41] Vgl. Ryte Wiki (2020): https://de.ryte.com/wiki/Natural_Language_Processing

[42] Vgl. Amazon Hilfe und Kundenservice: https://www.amazon.de/gp/help/customer/display.html?nodeId=201602230#GUID-1CDA0A16-3D5A-47C1-9DD8-FDEDB10381A3__SECTION2

dass Emotionen, die Stimmung oder der Dialekt Anweisungen beeinflussen können und diese oft nicht eindeutig und präzise sind.[43]

NLP klassifiziert sich in folgenden zwei Teilbereichen:	
Natural Language Understanding (NLU)	Nach der Spracherkennung, folgt die Zuordnung der Bedeutung des Sprachbefehls durch NLU.[44] Unstrukturierte Daten werden in für eine Maschine verstehende Form umgewandelt. Die Technologie ist in der Lage, bestimmte Elemente, wie Personen, Zeitpunkte, Orte zu erkennen und den Sinn hinter dem Sprachbefehl zu verstehen. Eingesetzt wird NLU in Chatbots und Sprachassistenten, um Nutzerabsichten zu erkennen und zielgerichtet zu antworten.[45]
Natural Language Generation (NLG)	Nach der Analyse und Verarbeitung der Nutzerabsicht, erfolgt die Sprachsynthese mit Hilfe der Technologie NLG. Diese ist verantwortlich für die Generierung von strukturierten Daten in natürliche Sprache. Mittels Algorithmen können lesefreundliche Texte auf Datenbasis generiert werden.[46] Einsatzgebiete von NLG sind z.B. Nachrichtenportale mit Artikeln mit vielen standardisierten Formulierungen, wie Berichterstattungen, Börsen- oder Sportnachrichten.[47]

Tabelle 3: Klassifizierung der Technologie NLP[48]

5.3. Technischer Dialogablauf

Nachfolgend wird ein technischer Dialogablauf am Beispiel des cloudbasierten Sprachassistenten Alexa beschrieben. Um die nachfolgenden Schritte besser zu verstehen, wird ein laufendes *Beispiel* verwendet, das in den einzelnen Schritten wiederholend aufgeführt wird. Das *Beispiel* bezieht sich auf die Interaktion mit einem Smart Speaker. Das Ziel des Anwenders besteht darin, das Wort „Sicherheit“ auf Englisch übersetzen zu lassen.

[43] Vgl. Ryte Wiki (2020): https://de.ryte.com/wiki/Natural_Language_Processing

[44] Vgl. P. Kabel (2020) S. 43

[45] Vgl. Retresco (2021): https://www.retresco.de/lexikon/natural-language-understanding/

[46] Vgl. A. Hoffmann (2019) S. 115

[47] Vgl. Retresco (2021): https://www.retresco.de/lexikon/natural-language-generation/

[48] Eigene Darstellung

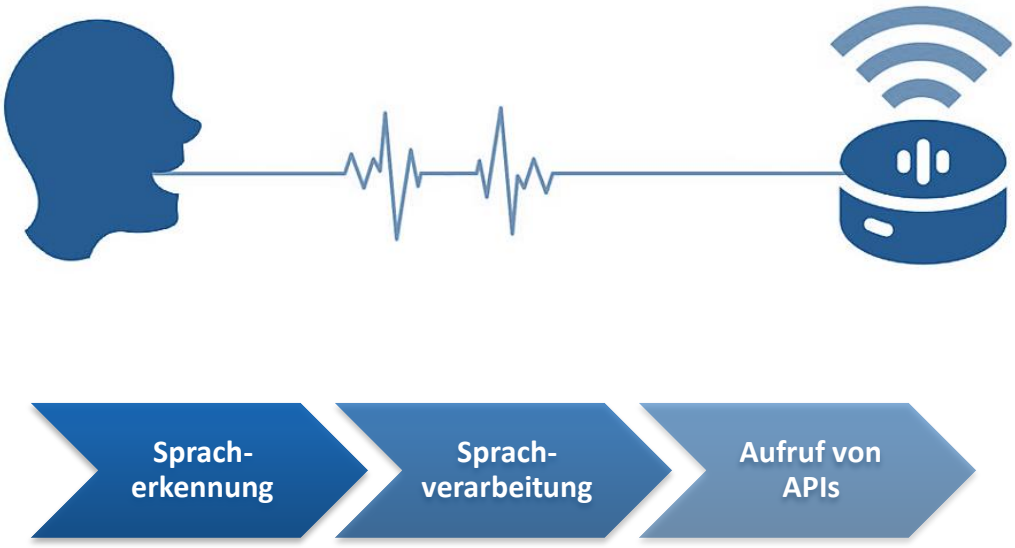

Abbildung 6: Dialogablauf zwischen Anwender und Sprachassistenten[49]

1. Aktivierung des Sprachassistenten

Im ersten Schritt erfolgt eine Aktivierung des Sprachassistenten durch das Aktivierungswort „Alexa“. Das System erkennt dieses Wort und löst eine Sprachaufzeichnung aus.[50]

Der Anwender teilt dem Smart Speaker per Sprachbefehl mit „Alexa, frage PONS, was heißt Sicherheit auf Englisch?“

2. Weiterleitung an die Cloud

Der Sprachbefehl wird über das Internet an die Cloud des Service-Anbieters gesendet, wo die Analyse des Sprachbefehls stattfindet. Die Cloud ist eine selbstlernende Software, die mit Hilfe von Spracheingaben stetig dazulernt.[51]

Der Sprachbefehl „Alexa, frage PONS, was heißt Sicherheit auf Englisch?“ wird an die Amazon Cloud „Amazon Voice Service“ gesendet.

3. Sprachverarbeitung in der Cloud

In der Cloud des jeweiligen Softwareanbieters erfolgt die Verarbeitung des Sprachbefehls mittels der Technologien ASR und NLP. Der Sprachbefehl des Anwenders wird

[49] In Anlehnung an: D. Pickell (14. Februar 2019): https://learn.g2.com/voice-assistant

[50] Vgl. M. Wendel (19. Mai 2019): https://www.homeandsmart.de/wie-funktioniert-alexa

[51] Vgl. M. Wendel (19. Mai 2019): Wie funktioniert Alexa? Bauweise und Funktionen einfach erklärt (homeandsmart.de)

von Hintergrundgeräuschen gefiltert und eine Worterkennung wird durchgeführt. Um den Inhalt zu verstehen, werden die Wörter analysiert, ob sie entscheidend für die Nutzerabsicht sind, auch „Intent“ genannt, oder nicht und entsprechend voneinander getrennt. Die extrahierte Nutzerabsicht kann über zwei Wege beantwortet werden:

3a. Beantwortung durch Zugriff auf Datenbank

Bei Sprachanfragen bezüglich Wetter, Rechenaufgaben oder Begriffserklärungen wird auf Datenbanken des Software-Anbieters zugegriffen.

Der Name der App „PONS“ wird erkannt. Dementsprechend wird nicht auf die Datenbank der Alexa-Cloud zugegriffen.

3b. Beantwortung mittels einer Voice-App

Bei Sprachbefehlen mit der Formulierung „[Aktivierungswort], frage [Name der App]…“ wird sich über Programmierschnittstellen, auch APIs genannt, auf jeweilige Voice-Apps bezogen.[52] Durch APIs können benötigte Informationen über verschiedenen Wissensdatenbanken beschafft werden.

Die Cloud „Amazon Voice Service“ schickt den Sprachbefehl an den Alexa-Skill „PONS“. Dort wird die Anfrage bearbeitet. Die Antwort wird anschließend zurück an die Cloud gesendet.

4. Sprachausgabe

Zum Schluss wird die Antwort in Textform in Sprache generiert mittels der Technologie Text-To-Speech und an das jeweilige Gerät zurückgeschickt. Die Antwort der Sprachanfrage wird in auditiver Form dem Anwender übermittelt.[53]

Der Smart Speaker teilt dem Anwender mit: „Das Wort ‚Sicherheit‘ heißt auf Englisch ‚safety‘

[52] Vgl. T. Hörner (2019): https://sprachassistenten.marketing/allgemein/wie-funktionieren-sprachassistenten/

[53] Vgl. M. Wendel (19. Mai 2019): Wie funktioniert Alexa? Bauweise und Funktionen einfach erklärt (homeandsmart.de)

5.4. Zusammenfassung

Die Herausforderung von Sprachassistenten liegt darin, dass die Nutzerabsicht erkannt, verstanden und ausgeführt wird. Hierfür werden Technologien, wie ASR und NLP mit den Teilbereichen NLU und NLG eingesetzt. Diese entwickeln sich mit hoher Geschwindigkeit weiter. In Zukunft ist zu erwarten, dass sich die Sprachverarbeitung deutlich optimieren wird.

6. Stärken und Schwächen von Sprachassistenten

In diesem Kapitel werden die Stärken und Schwächen von Sprachassistenten näher beleuchtet.

6.1. Stärken von Sprachassistenten

Laut des Bundesverbands Digitale Wirtschaft (BVDW) ist die Sprache nicht nur schneller als das Tippen am Smartphone, sondern auch die natürlichste Art und Weise der Kommunikation.[54] Daraus erschließt sich ein hohes Potenzial von Sprachassistenten. Das Augenmerk richtet sich vor allem auf die Bewegungsfreiheit, intuitive und schnelle Interaktion sowie auf die kontinuierliche Einsatzbereitschaft.

6.1.1. Bewegungsfreiheit und Ortsunabhängigkeit

Die Bedienung von Sprachassistenten erfolgt berührungsfrei und ortsunabhängig, weder ein Blick noch eine Berührung wird benötigt. Anwender können sich bei Sprachanwendungen frei bewegen und sich parallel mit anderen Aufgaben beschäftigen. Beachtet werden muss lediglich, dass sich die Mikrofone des Sprachassistenten in Hörweite befinden.[55]

Ein Sprachassistent eignet sich daher besonders für Tätigkeiten, bei denen der Fokus auf Hände und Augen gerichtet ist, wie zum Beispiel beim Autofahren oder beim Durchführen von Reparaturen. Anwender können sich auf das Wesentliche fokussieren und ihren Blick auf den Weg richten bzw. auf den zu reparierenden Gegenstand.

6.1.2. Intuitiv

Eine weitere Stärke von Sprachassistenten liegt darin, dass die Interaktion mittels natürlicher Sprache erfolgt. Dafür muss weder eine weitere Tätigkeit erlernt werden noch wird Vorwissen benötigt.[56] Davon können vor allem Menschen profitieren, die mit neuer Technik kaum Berührungspunkte haben. Demzufolge sind Sprachassistenten für die Zugänglichkeit von Technik bedeutend.

---

[54] Vgl. Bundesverband Digitale Wirtschaft (BVDW) e.V. (April 2019): https://www.bvdw.org/fileadmin/bvdw/upload/publikationen/smart_home/Sprachassistenten_im_SmartHome.pdf

[55] Vgl. T. Hörner (2019) S. 39

[56] Vgl. T. Hörner (2019) S. 5

6.1.3. Schnelligkeit

Die Eingabe über die Sprache ist wesentlicher schneller als eine Texteingabe. Das geht aus der im Jahr 2016 durchgeführten Studie „Speech Is 3x Faster than Typing for English and Mandarin Text Entry on Mobile Devices“ der Standford Universität und dem chinesischen Unternehmen Baidu hervor. An der Studie nahmen insgesamt 38 Personen im Alter von 19 bis 32 Jahren teil.

Hierbei wurde die Geschwindigkeit von Tastatureingaben mit Spracheingaben bei mobilen Endgeräten verglichen. Die Studie wurde in den Sprachen Englisch und Mandarin-Chinesisch durchgeführt. Es stellte sich heraus, dass die Spracherkennungssoftware dreimal schneller den Text auf mobilen Endgeräten erzeugte und zugleich 20 Prozent weniger fehleranfällig als Tastatureingaben war.[57]

6.1.4. Sprachsuchen

Besonders Sprachsuchen bieten eine Reihe von Vorteilen für Anwender. „Voice Search“ bezeichnet den Vorgang, wenn Anwender für die Sucheingabe die natürliche Stimme verwenden.[58] Sprachsuchen sind zeitsparend und für Anwender komfortabler als Texteingaben. Besonders Personen, die Schwierigkeiten mit dem Umgang einer Tastatur haben, können von Spracheingaben und –suchen profitieren. Dazu zählen beispielsweise ältere Anwender, Personen mit motorischen Beeinträchtigungen oder Personen mit einer Rechtschreibschwäche.[59]

6.1.5. Gezielte Sprachausgabe

Die oft komplexen Navigationsstrukturen von Webseiten stellen vielen Personen vor Herausforderungen. Oftmals lassen sich erst nach mehreren Klicks und langem Suchen relevante Informationen finden. Dieses Problem umgehen Anwender von Sprachassistenten. Mit einer Spracheingabe erledigt die Software diesen Vorgang. Somit gelangen Anwender schnell und gezielt an gewünschte Informationen.[60]

---

[57] Vgl. J. Landay et al. (23. August 2016): https://hci.stanford.edu/research/speech/paper/speech_paper.pdf

[58] Vgl. P. Kabel (2020) S. 201

[59] Vgl. Ryte Wiki (2020): https://de.ryte.com/wiki/Voice_Search

[60] Vgl. P. Kabel (2020) S. 10 f.

6.1.6. Verfügbarkeit

Darüber hinaus sind Sprachassistenten jederzeit verfügbar, denn sie sind standardmäßig eingeschaltet und befinden sich in Wartestellung bis das Aktivierungswort ausgesprochen wird.[61] Folglich können Anwender unmittelbar mit Sprachassistenten kommunizieren.

6.1.7. Intelligenz

Die amerikanische Agentur „Stone Temple“ ging der Frage nach, wie hoch die Intelligenz verschiedener Sprachassistenten ist. Insgesamt wurden 5.000 gleiche Fragen zum Allgemeinwissen an Sprachassistenten, wie Google Assistant, Cortana, und Alexa gestellt.

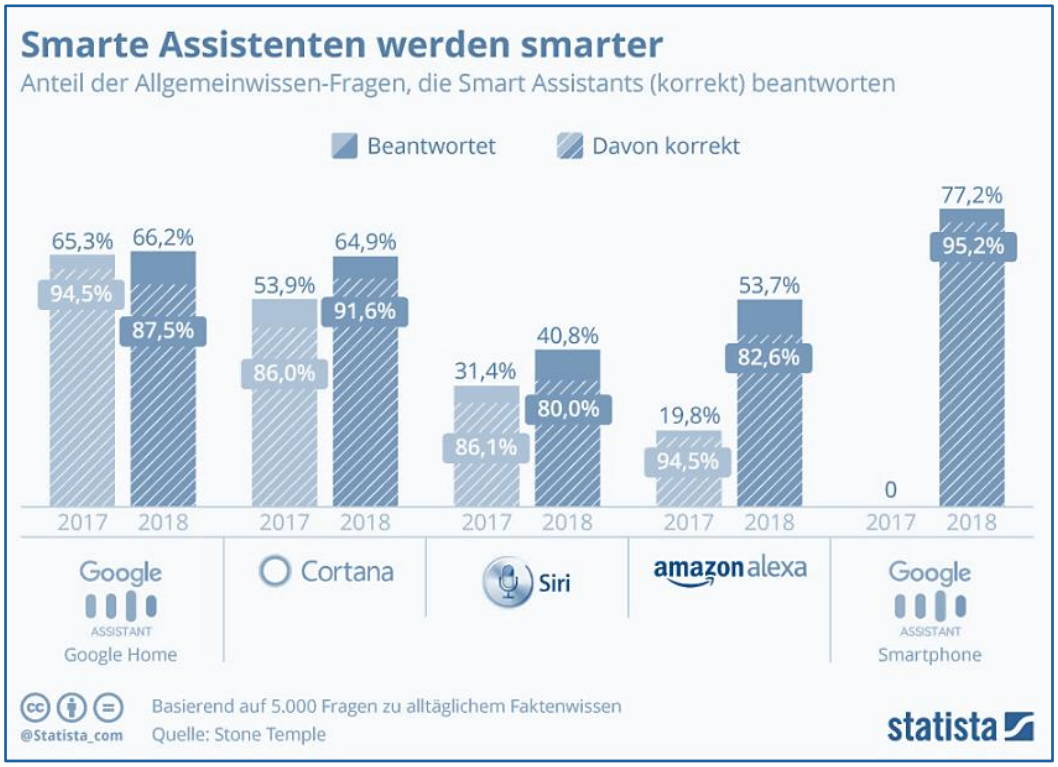

Abbildung 7: Steigende Intelligenz von Sprachassistenten[62]

Der Sieger des Tests war Google Assistant auf dem Smartphone. Die Antwortquote lag bei 77,2 Prozent. Davon waren insgesamt 95,2 Prozent der Antworten korrekt. Der Verlierer ist Siri von Apple mit einer Antwortquote von 40,8 Prozent und 80 Prozent

[61] Vgl. U. Albrecht (2020) S.10

[62] Vgl. M. Brandt (25. April 2018): https://de.statista.com/infografik/9615/allgemeinwissen-von-smart-assistants/

richtigen Antworten. Zusammenfassend stieg die Intelligenzquote bei jedem Sprachassistenten innerhalb eines Jahres an.

6.2. Schwächen von Sprachassistenten

Eine neue Technologie bringt auch Herausforderungen mit sich. Im Folgenden werden Risiken von Sprachassistenten erläutert.

6.2.1. Fehleranfälligkeit

Eine Schwäche von Sprachassistenten liegt darin, dass die Intention des Anwenders nicht verstanden wird. Der Anwender erhält kein Feedback und wird nicht über die Ursache des Problems informiert. Eine Studie beschäftigte sich mit den Fehlerquellen von sprachlichen Benutzerschnittstellen. In der veröffentlichten Studie im Jahr 2018 „Pattern for How Users Overcome Obstacles in Voice User Interfaces" wurden die vier häufigsten Fehlerquellen analysiert.

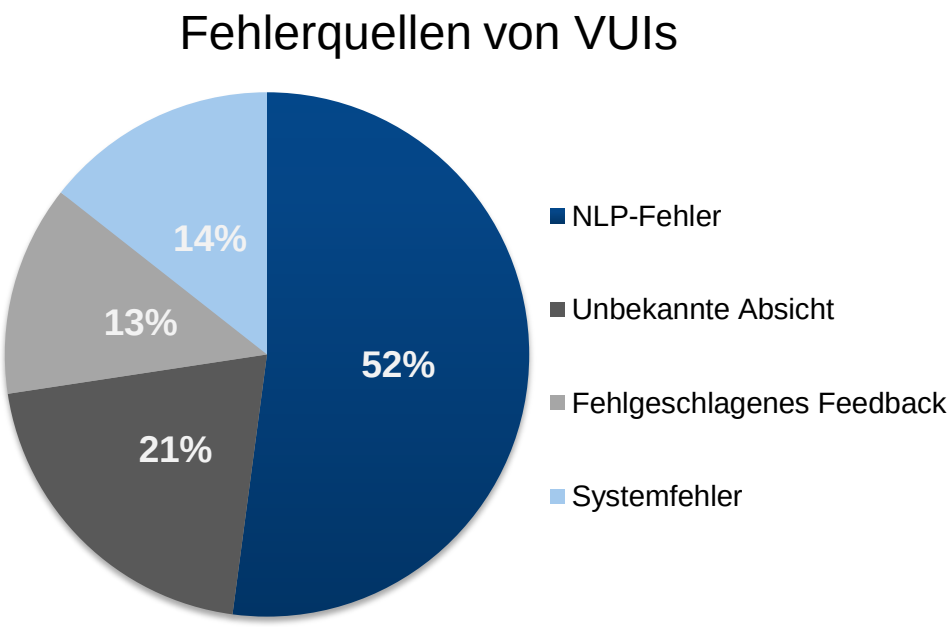

Abbildung 8: Häufige Fehlerquellen eines VUIs[63]

Die größte Fehlerquelle mit 52,1 Prozent sind Fehler, die mit der Technologie Natural Language Processing in Verbindung gebracht werden. Beispielsweise wird die Spracheingabe falsch verstanden und dementsprechend wird eine falsche Absicht zugeteilt.

[63] In Anlehnung an: C. Myers et al. (April 2018): https://www.researchgate.net/publication/324670825_Patterns_for_How_Users_Overcome_Obstacles_in_Voice_User_Interfaces

Danach folgt eine unbekannte Absicht mit 20,5 Prozent. Wenn Probanden das VUI-System nicht kennen kann es dazu führen, dass entweder die Äußerung von der Technologie nicht verstanden wird oder die Äußerung vom VUI-System nicht unterstützt wird. Das Problem liegt hierbei, dass unerfahrene Anwender nicht wissen, wie die Befehle artikuliert werden oder welche Befehle vom System ausführbar sind.

14,4 Prozent stellten Systemfehler als Fehlerquelle von VUIs dar. Ein Systemfehler entsteht bei einem Fehler der Architektur eines VUI-Systems. Beispielsweise mussten Anwender mehrmals eine Intention ausführen lassen, um das gewünschte Ergebnis zu erzielen. Zuletzt gab es noch mit 13 Prozent die Fehlerkategorie des Feedbacks. Zum Teil gab es kein oder mangelndes Feedback, wenn ein Fehler aufgetreten ist. Aufgrund fehlender Informationen sind die Anwender nicht an das gewünschte Ziel gelangen.[64]

6.2.2. Informationseinschränkung

Eine weitere Schwäche bei der Benutzung von Sprachassistenten, wie z.B. bei Smart Speaker, ist das Fehlen eines Displays und somit auch der strukturelle Aufbau, wie z.B. eine Inhalts- oder Navigationsstruktur. Ausschließlich auditive und keine visuellen Informationen werden vermittelt. Sprachassistenten eignen sich daher für einfach zu vermittelnde Informationen, wie z.B. für Antworten auf konkreten Fragen. Hier rückt der Begriff „Multimodalität“ in den Vordergrund. Immer mehr Sprachassistenten kombinieren einen Bildschirm, wie z.B. ein Smart Display, der in diesem Kapitel 7.3. beschrieben wird.

6.2.3. Umgebungsgeräusche

Ein weiteres Problem ist die Gefahr, dass aufgrund lauter Umgebungs- und Hintergrundgeräusche, die Spracheingabe akustisch nicht verstanden wird. Die Sprachassistenten sind zwar mit einer Fernfeld-Technologie sowie mit Fernfeld-Mikrofonen ausgestattet, dennoch kann es passieren, dass Spracheingaben von ungewollten Geräuschen, wie sprechende Personen, nicht trennbar sind.[65]

6.2.4. Öffentlichkeit

Des Weiteren muss bedacht werden, dass bei Sprachanwendungen Anwender persönliche Informationen in den öffentlichen Raum laut aussprechen. Dabei besteht die Ge-

[64] Vgl. C. Myers et al. (April 2018): https://www.researchgate.net/publication/324670825_Patterns_for_How_Users_Overcome_Obstacles_in_Voice_User_Interfaces

[65] Vgl. P. Kabel (2020) S. 160

fahr, dass die Privatsphäre angegriffen wird. Die Digitalagentur Beyto, die auf Sprache und Sprachassistenten spezialisiert ist, veröffentlichte eine Studie, bei der 2.042 Personen im März 2020 über die Nutzung von Smart Speaker befragt wurden. Daraus ging hervor, dass 60 Prozent der Befragten der Ansicht waren, dass die Privatsphäre ein großes Risiko darstellt. Dies sei auch der Grund, warum viele Personen keine Smart Speaker besitzen möchten.[66]

6.2.5. Datenschutz

Ein weiteres Problem stellt der Datenschutz dar. Durch fehlende Transparenz sind sich Anwender unsicher, ob ihre Daten sicher sind. Dies zeigt eine repräsentative Studie von Microsoft, die im Jahr 2019 durchgeführt wurde. Die Studie beschäftigte sich mit der Frage, wie Anwender bei der Nutzung von Sprachassistenten über Datensicherheit denken.

52 Prozent der Befragten haben großes Bedenken, dass ihre persönlichen Daten nicht sicher sind. 41 Prozent sind der Meinung, dass Sprachassistenten passiv mithören und Gespräche aufzeichnen und 36 Prozent möchten nicht, dass persönliche Daten verwendet werden.[67] Dieses Bedenken ist nicht unbegründet. Aus dem Informationsschreiben „Miniaturisierung der Informations- und Kommunikationstechnik“ der Bundesbehörde für Datenschutz und Informationsfreiheit aus dem Jahr 2017 geht folgendes hervor:

Aus Sicht des Datenschutzes ist es problematisch, dass die Stimmen der Anwender sowie jegliche Hintergrundgeräusche kontinuierlich aufgenommen werden, da Sprachassistenten standardmäßig eingeschalten sind. Es sei unklar, wo und wie die aufgezeichneten Daten verwendet und gespeichert werden. Aus dem Dokument geht hervor, dass Hersteller von Assistenzsystemen gegenüber Anwender transparent sein sollen und sich offen zur Datenspeicherung und deren Verwendung äußern müssen.[68]

Der Hersteller Amazon geht beispielsweise mit Datenschutz und -verarbeitung so vor, dass Anwender die Historie der Sprachdialoge in der Alexa-App kontrollieren und lö-

---

[66] Vgl. C. Konopka, M. Metzen (März 2020): https://mcusercontent.com/cab8c0a7e9a3fef590fd49ce1/files/3cf2966b-5b46-4f79-a597-80fcdb04fcff/Beyto_Smart_Speaker_Studie_2020_Studienreport.pdf

[67] C. Olson, K. Kemery (2019): bingads_2019_voicereport.pdf (azureedge.net)

[68] Vgl. (07.03.2017): Die Bundesbeauftragte für den Datenschutz und die Informationsfreiheit (2021): https://www.bfdi.bund.de/DE/Datenschutz/Themen/Technische_Anwendungen/TechnischeAnwendungenArtikel/Miniaturisierung.html

schen können. Des Weiteren kann eingestellt werden, dass Suchanfragen auf jeweilige Alexa-Geräte nach einer bestimmten Zeit automatisch gelöscht werden.[69]

Bei einem Interview hat sich der Country Manager Herr Dr. Berger zum Datenschutz und Privatsphäre von Alexa geäußert. Auf die Frage, ob jedes Gespräch an die Cloud von Amazon gesendet wird, antwortete er, dass erst nach Aussprache und Identifikation des Aktivierungswortes, das mit Hilfe eines Sprachmusters erkannt wird, Audiodateien an die Cloud gesendet werden.

Allerdings kann die Gefahr bestehen, dass das System ein ähnlich klingendes Wort, wie z.B. Alexander, versteht. An der verbesserten Erkennung des Aktivierungswortes wird derzeit gearbeitet. Herr Dr. Berger ergänzt hinsichtlich der Datenverarbeitung, dass Spracheingaben zum Training der Algorithmen benutzt werden. Das Ziel liegt hierbei, dass Alexa in Zukunft Sprachbefehle noch besser verstehen und darauf reagieren kann. Anwender haben die Möglichkeit, die Datenverwertung für Trainingszwecke abzulehnen.[70]

Der Wissenschaftliche Dienst des Bundestags zieht in einem Gutachten bezüglich Datensicherheit des Herstellers Amazon das Fazit, dass Amazon seiner Pflicht zu Informationsvermittlung gemäß Art. 13 I und Art. 14 I DSGVO in hinreichendem Maß nachgekommen sein dürfte.[71]

[69] Vgl. M. Wendel (07. Dezember 2020): Alexa Datenschutz: Wie viel Sicherheit bietet Amazon Alexa? (homeandsmart.de)

[70] Vgl. D. Wulf (18. Dezember 2018): https://www.homeandsmart.de/alexa-datenschutz-interview-amazon

[71] Deutscher Bundestag (29. Mai 2019): Microsoft Word - WD 10-032-19 S Alexa.docx (bundestag.de)

6.3. Zusammenfassung

In der anknüpfenden Tabelle werden die Stärken und Schwächen von Sprachassistenten zusammengefasst.

Stärken:	**Schwächen:**
1. Bewegungsfreiheit: ▪ „Hands-free“ ▪ „Eyes-free“ ▪ Ortsunabhängig **2.** Intuitive Spracheingaben **3.** Geschwindigkeit ▪ Spracheingaben dreimal schneller als Texteingaben ▪ Spracheingaben um 20 Prozent weniger fehleranfällig als Texteingaben **4.** Sprachsuchen **5.** Gezielte Sprachausgabe **6.** Verfügbarkeit **7.** Steigende Intelligenz	**1.** Fehlerquellen: ▪ NLP-Fehler ▪ Unbekannte Absicht ▪ Fehlgeschlagenes Feedback ▪ Systemfehler **2.** Informationseinschränkung **3.** Zu laute Umgebungs- oder Hintergrundgeräusche **4.** Fehlende Privatsphäre **5.** Datenschutz: ▪ Nicht transparente Gestaltung bezüglich der Datenverarbeitung ▪ Gefahr von unbeabsichtigten Aufzeichnungen

Tabelle 4: Gegenüberstellung von Stärken und Schwächen von Sprachassistenten[72]

Auf der einen Seite sprechen deutliche Argumente für Sprachassistenten, wie intuitive Spracheingaben und Sprachsuchen, die nicht nur schneller als Texteingaben sind, sondern auch weniger fehleranfällig. Auf der anderen Seite müssen, ähnlich zur Nutzung von Smartphones, persönliche Informationen von Anwendern herausgegeben werden, um die Vorteile von Sprachassistenten in vollem Umfang nutzen zu können.

[72] Eigene Darstellung

7. Funktionsanalyse und Produktbeschrei- bung des Sprachassistenten Alexa

Dieses Kapitel beschäftigt sich mit einer Produkt- und Funktionsanalyse am Beispiel des Sprachassistenten Alexa. Das erste Unterkapitel erläutert den grundlegenden Funktionsumfang von Alexa. Im Anschluss wird näher auf die Drittanbieter-Apps namens „Skills“ eingegangen. Im Weiteren wird das Smart Display „Echo Show 10“ von Amazon vorgestellt und weitere Funktionen erläutert.

7.1. Grundlegender Funktionsumfang

Voraussetzung für die Nutzung des cloudbasierten Sprachassistenten Alexa, ist die Herstellung einer Internetverbindung, denn die Ausführung des Sprachbefehls erfolgt auf Servern der Amazon-Cloud. Zudem sollte die Alexa-App am Smartphone oder Tablet heruntergeladen werden, um das jeweilige Gerät einzurichten und individuelle Einstellungen, wie z.B. Datenverarbeitung, festzulegen. Die App ist für iOS- und Android-Geräte zum Download kostenlos verfügbar.[73]

Der Sprachassistent Alexa stellt einen Umfang an grundlegenden Funktionalitäten zur Verfügung, die beispielhaft in der folgenden Tabelle aufgelistet werden:

Informationsabruf:
▪ Uhrzeit ▪ Aktuelle Nachrichten ▪ Örtliche Verkehrs- und Wetterinformationen ▪ Örtliche Suche, wie Informationen zu Öffnungszeiten oder Adressen ▪ Beantwortung von Fragen bezüglich des Allgemeinwissens
Unterhaltung:
▪ Abspielen von Musik, Radio und Hörbüchern ▪ Abspielen von Fotos und Filmen
Organisation:

[73] Vgl. M. Wendel (11. Januar 2021): https://www.homeandsmart.de/how-to-amazon-alexa-app-effektiv-nutzen

▪ Stellung von Erinnerung, Timer und Wecker ▪ Verknüpfung mit Kalender, um Ereignisse abzufragen oder hinzuzufügen ▪ Erstellung von Routinen, um mit einem Sprachbefehl mehrere Aktionen auszuführen ▪ Erstellung von Listen, wie z.B. Einkaufslisten oder To-Do-Listen ▪ Steuerung von Smart-Home-Geräten
Kommunikation:
▪ Tätigen und Entgegennahme von Anrufen ▪ Versenden und Empfangen von Nachrichten

Tabelle 5: Grundlegende Funktionen des Sprachassistenten Alexa[74]

7.2. Alexa-Skills

Voice-Apps sind Programme bzw. Funktionalitäten, die von Drittanbietern erstellt werden und sind vergleichbar mit mobilen Apps. Die Anbieter von Sprachassistenzsystemen stellen die Grundfunktionalitäten zur Verfügung und Drittanbieter integrieren eigene Angebote durch Voice-Apps.[75] Das Unternehmen Amazon verfolgt bei der Hardware eine offene Lizenzpolitik. Dies ermöglicht, dass unterschiedliche Endgeräte-Hersteller den Sprachassistenten Alexa nutzen, um ihre eigenen Produkte mit intelligenter Sprachsteuerung zu versehen und damit einen zusätzlichen Kanal für Verbraucher anzubieten.[76]

Der Funktionsumfang mit dem Sprachassistenten Alexa lässt sich durch die Aktivierung von sogenannten „Skills“ erweitern. Ein „Skill“ ist eine Programmierung auf dem Sprachassistenzsystem und wird mit dem Wort „Alexa“ aktiviert. Der Name des „Skills“ wird von den Programmierern bestimmt. Mit einem „Skill“ können Sprachbefehle erweitert werden und dadurch die Kommunikation von Alexa personalisieren.[77]

Als Beispiel für einen „Skill“ ist „Mercedes-me“ des Automobilherstellers Daimler zu nennen. Mit der Aktivierung dieses „Skills“ lassen sich Fahrzeuge der Automarke Mer-

[74] Vgl. Amazon (2021): https://www.amazon.de/b?ie=UTF8&node=12775495031

[75] Vgl. R. Kreutzer, D. Vousoghi (2020) S.44 f.

[76] Vgl. Deloitte: https://www2.deloitte.com/de/de/pages/consumer-business/articles/sprachassistent.html

[77] Vgl. T. Hörner (2019): https://sprachassistenten.marketing/grundlagen/voice-apps-was-ist-das/

cedes per Sprachbefehl fernsteuern. Dadurch kann sich der Anwender über den Fahrzeugstatus informieren, aber auch effektiv Dienste starten bzw. beenden, wie zum Beispiel ferngesteuert die Türen schließen oder die Standheizung starten.[78] Dabei sendet der „Skill“ Fernbefehle an das eigene Fahrzeug. Benötigt wird ein aktives Mercedes-me-Konto. Der „Skill“ wird mit dem Sprachbefehl „Alexa, starte Mercedes“ aktiv. Danach folgt ein Kommando, wie zum Beispiel „Alexa, frage Mercedes, wann ist der nächste Service-Termin?“ oder „Alexa, sage Mercedes, starte die Standheizung.“[79]

7.3. Produktbeschreibung eines Smart Displays

Im Folgenden wird das Produkt „Echo Show 10“ beschrieben, das im September 2020 von Amazon präsentiert wurde. Es handelt sich hierbei um das Smart Display der dritten Generation. Der Grund für die Wahl des Produkts liegt darin, dass das Smart Display eine sprachbasierte mit einer visuellen Benutzerschnittstelle kombiniert und sich daraus weitere nützliche Funktionen ergeben.

7.3.1. Technische Details

In der folgenden Abbildung ist der Smart Display „Echo Show 10“ in der Vorder- und Hinteransicht dargestellt mit einer anschließenden Tabelle, die technische Details des Smart Displays auflistet.

Abbildung 9: Gerätebeschreibung des Smart Displays Echo Show 10[80]

[78] Vgl. R. Kreutzer, D. Vousoghi (2020) S. 15

[79] Vgl. Amazon (2021): https://www.amazon.de/Mercedes-Benz-Mercedes-me/dp/B075435C1Y

[80] Vgl. Amazon (2021): https://www.amazon.de/der-neue-echo-show-10-hochauflosendes-smart-display-mit-bewegungsfunktion-und-alexa-anthrazit/dp/B084P3KP2S

Technische Details des Smart Displays „Echo Show 10“:	
Bewegungsfunktion	Motor mit 175-Grad-Drehradius
Display	10,1-Zoll-Touchscreen
Farben	Anthrazit und Weiß
Gewicht	2560 Gramm
Größe	251 x 230 x 172 Millimeter
Kamera	13 Megapixel
Material	Recyceltem Textilien bzw. Aluminium
Prozessoren	Hauptprozessor: MediaTek 8183 Nebenprozessor: AZ1 Neural Edge
Sensor	RGB-Umgebungslichtsensor
Soundsystem	76-Millimeter-Woofer und 25-Millimeter-Hochtonlautsprecher
Smart Home-Hub	ZigBee
Sprachassistenz	Alexa
Sprachen	Deutsch, Englisch, Französisch, Italienisch und Spanisch

Tabelle 6: Technische Details des Smart Displays „Echo Show 10“[81]

7.3.2. Weitere Funktionen von Smart Displays

Zu dem grundlegenden Funktionsumfang, der im Abschnitt erläutert wurde, erschließen sich durch die Erweiterung der visuellen Schnittstelle weitere nützliche Funktionen im Bereich der Technischen Kommunikation, die im Folgenden näher beschrieben werden.

Darstellung visueller Inhalte:

Neben der Sprachausgabe werden zusätzliche visuelle Inhalte, wie z.B. unterstützende Navigationselemente auf einem Bildschirm dargestellt, der als vertraute Oberfläche

[81] Vgl. Amazon (2021): https://www.amazon.de/der-neue-echo-show-10-hochauflosendes-smart-display-mit-bewegungsfunktion-und-alexa-anthrazit/dp/B084P3KP2S

gilt.[82] Durch den integrierten Motor ist der Smart Display in der Lage, den Bildschirm und die Mikrofone in Richtung Anwender zu richten. Die automatisierte Schwenkfunktion hat den Vorteil, dass Anwender trotz Bewegungen automatisch im Kamerafokus bleiben und der Bildschirm optimal positioniert ist.[83]

Der bewegliche Bildschirm ist vor allem vorteilhaft, wenn Anwender parallel mit Tätigkeiten beschäftigt sind und die Hände frei bleiben müssen, z.B. beim Ausführen von Handlungsschritten bei technischen Geräten.

Bewegte Bilder, wie z.B. Animationen, Simulationen oder Videos können Inhalte besser an Anwender vermitteln als eine Text-Bild-Kombination. Um Handlungen multimedial darzustellen, eignen sich Videos oder Tutorials.[84] Diese können über einen Smart Display abgespielt werden. Über Sprachsteuerung kann der Anwender nach jedem auszuführenden Handlungsschritt das Video pausieren. In der Zeit des pausierten Videos erfolgt die direkte Ausführung des Handlungsschrittes. Danach kann das Video oder das Tutorial fortgesetzt werden. Der Bildschirm ist dabei immer zum Anwender gerichtet, sodass dieser sich frei bewegen kann und jederzeit das Video bzw. das Tutorial im Blick hat.

Hinsichtlich des Datenschutzes findet die Datenverarbeitung lokal statt. Außerdem können die Mikrofone aufgeschaltet werden und die Kamera durch die integrierte Abdeckung verdeckt werden. Die Bewegungsfunktion kann per Sprachbefehl am Smart Display oder über die Alexa App deaktiviert werden.[85]

Steuerungszentrale im Smart Home:

Die Steuerungszentrale ZigBee-Hub ist im Smart Display „Echo Show 10“ integriert. ZigBee gilt als Smart Home-Funkstandard und stellt eine Datenübertragung zwischen intelligenten Geräten her. Der Datenaustausch verläuft über ein über Funk eigenständiges Netzwerk. Die Vorteile von ZigBee liegen zum einen in der hohen Energieeffizient und zum anderen in der flexiblen Einsetzbarkeit. Sobald Geräte im Netzwerk eingebunden sind, läuft die Hausautomation eigenständig.[86] Über das Smart Display lassen sich kompatible Geräte vernetzen, wie z.B. Lampen, Thermostate oder Tür-

[82] Vgl. T. Kahle, D. Meißner (2020) S. xii

[83] Vgl. M. Wendel (14. April 2021): https://www.homeandsmart.de/echo-show-10-test

[84] Vgl. D. Juhl (2015) S. 350 f.

[85] Vgl. Amazon (2021): https://www.amazon.de/der-neue-echo-show-10-hochauflosendes-smart-display-mit-bewegungsfunktion-und-alexa-anthrazit/dp/B084P3KP2S

[86] Vgl. M. Wendel (07. Dezember 2020): https://www.homeandsmart.de/zigbee-funkprotokoll-hausautomation

schlösser. Über das angezeigte Dashboard kann der Stromverbrauchsmodus angezeigt werden.[87]

7.4. Zusammenfassung

Der cloudbasierte Sprachassistent Alexa bietet Anwendern viele nützliche Funktionen rund um Bereiche wie Informationsabruf, Unterhaltung, Organisation sowie Kommunikation. Darüber hinaus kann der Funktionsumfang erweitert werden mit Aktivierung von „Skills“. Diese ermöglichen eine individualisierte Nutzung mit dem Sprachassistenten. Durch die Erweiterung eines Bildschirms wird neben dem Hörsinn, auch der Sehsinn beansprucht. Daraus entstehen nützliche Funktionen und Vorteile für Anwender.

[87] Vgl. M. Baumann (11. November 2019): Amazon Echo Show: Alle Smart Display Modelle im Test-Überblick (homeandsmart.de)

8. Anwendungsfälle von Sprachassistenten in der Technischen Kommunikation

Dieses Kapitel befasst sich mit Anwendungsfällen in der Technischen Kommunikation, in denen Sprachassistenten nützliche Funktionen für Anwender bieten. Dabei wird Bezug genommen auf die Bereiche Fremdsprachenübersetzung, Technische Dokumentation und Kundenservice.

8.1. Fremdsprachenübersetzung

Sprachassistenten sind in der Lage, in Echtzeit Gespräche zu übersetzen. Dies ist vor allem bei der Nutzung von Smart Displays sinnvoll. Die Ausgangssprache sowie die übersetzte Zielsprache werden laut vorlesen und zugleich angezeigt. Beispielsweise bietet die Live-Übersetzungsfunktion Alexa und Google Assistant an, die per Zuruf aktiviert wird. Die Sprachassistenten fungieren dabei als Echtzeit-Übersetzer zwischen den Personen, die jeweils unterschiedliche Sprachen sprechen.

Bei Alexa findet eine automatische Erkennung der Sprache statt. Realisiert wird dies mit Hilfe eines Spracherkennungssystems und eines Text-To-Speech-Systems. Aktuell bietet Alexa eine Übersetzung zwischen Englisch und einer der folgenden Sprachen an: Deutsch, Französisch, Hindi, Italienisch, Portugiesisch oder Spanisch.[88] Deutlich mehr zu übersetzende Sprachen, stellt der Google Assistant dar. Bei der Übersetzungsfunktion am Smartphone und am Tablet kann in 44 Sprachen übersetzt werden. Beispielsweise wird mit dem Sprachbefehl „Ok Google, chinesischer Dolmetscher“ zwischen Deutsch und Chinesisch übersetzt.[89]

Die Vorteile bei dieser Funktion liegen in der unmittelbaren und zeitersparenden Übersetzung. Ebenso wird ein barrierefreier Zugang zu Wissen und Informationen gewährt. Die Übersetzungsfunktion bietet Vorteile bei Gesprächen in international tätigen Unternehmen, aber auch in privaten Situationen, in denen Personen sich aufgrund von Sprachbarrieren nicht verständigen können. Allerdings sollte beachtet werden, dass bei

[88] Vgl. I. Campbell (14. Dezember 2020): https://www.theverge.com/2020/12/14/22174561/amazon-alexa-live-translation-google-assitant-interpreter-mode

[89] Vgl. Google Assistant-Hilfe (2021): https://support.google.com/assistant/answer/9234753?hl=de

der maschinellen Übersetzung Übersetzungsfehler entstehen können, insbesondere bei langen Sätzen mit komplexen Satzstrukturen.[90]

Die Schwierigkeit der maschinellen Übersetzung liegt im Verständnis der Zusammenhänge von Sätzen. Auch Mehrdeutigkeiten stellen ein Problem dar. Wenn die Übersetzungsqualität einen hohen Stellenwert hat, wie beispielsweise im politischen und juristischen Umfeld, sollte die Übersetzung nicht die Maschine übernehmen, sondern professionelle Übersetzer.[91] Dies gilt auch im Bereich der Technischen Dokumentation. Hierbei besteht die Gefahr von Übersetzungsfehlern, die zu Fehlanwendungen führen können und dadurch kein bestimmungsgemäßer Gebrauch gewährleistet werden kann.

Das Erlernen einer Fremdsprache bietet einen nützlichen Anwendungsfall bezüglich Sprachassistenten in der Fremdsprachenübersetzung an. Als Beispiel kann der Sprachen-Lern-Lautsprecher „Lily“ aufgeführt werden. Das Pilotprojekt des US-Unternehmens „Maybe“ wird auf der Crowdsourcing-Plattform Indiegogo angeboten.[92] Die interaktive Konversation mit dem Lautsprecher und der Nutzung der dazugehörigen App lässt sich in Echtzeit-Übersetzung ein akustisches sowie individuelles Sprachtraining zusammenstellen. Neben der individuellen Lerngeschwindigkeit und dem Wegfall von Lehrbüchern, kann verstärkt eine richtige Aussprache der Vokabeln gewährleistet werden. Letzteres ist besonders im Chinesischen von hoher Relevanz, da bei der Aussprache die Tonhöhe die Bedeutung der Wörter verändert. Neben der Grundsprache Englisch kann der Lautsprecher vom Chinesischen ins Deutsche übersetzen.[93]

8.2. Technische Dokumentation

„Gute, verständliche und vor allem anwendbare Anleitungen sind ein wichtiges Glied in der Schnittstelle Mensch-Maschine, ermöglichen dem Anwender sich selbstständig mit dem Gerät vertraut zu machen, die Bedienung zu erlernen und Nutzen aus der Technik zu ziehen.“[94]

[90] Vgl. M. Olschewski (07. Februar 2019): https://www.homeandsmart.de/google-home-dolmetscherfunktion

[91] Vgl. A. Schmalz (2019) S. 196

[92] Vgl. Indiegogo (2021): https://www.indiegogo.com/projects/lily-the-1st-smart-speaker-that-teaches-chinese#/

[93] Vgl. P. Kabel (2020) S. 21 f.

[94] D. Juhl (2015) S. VII

Mit der Digitalisierung werden Bedienungsanleitungen in digitaler Form zunehmend wichtiger. Laut der Gesellschaft für Technische Kommunikation eröffnen sich, im Vergleich zur gedruckten Version, bei Bedienungsanleitungen in digitaler Form zahlreiche Vorteile für Anwender:

- Direkter und schneller Zugriff auf mobilen Endgeräten
- Jederzeit aktualisierbar
- Wechsel von verfügbaren Sprachen
- Such- und Vorlesefunktion
- Einbindung von multimedialen Inhalten (Links auf weiterführende Informationen oder Tutorials)
- Unterstützung von Anwendungen von Augmented-Reality-Brillen oder Trainingsmodellen in Virtual Reality[95]

Hinzu kommt, dass digitale Bedienungsanleitungen Spracheingaben und –ausgaben ermöglichen. Daraus entstehen weitere Vorteile, wie die Gewährleistung einer barrierefreien Nutzung und eine schnelle und intuitive Bedienung. Anwender können über Spracheingaben Begriffe suchen und über Sprachausgaben die Vorlesefunktion nutzen:

- **Sprachsuche:** Wie in Kapitel 6.1.3 beschrieben, sind Spracheingaben deutlich schneller als Texteingaben. Anstatt Begriffe über Texteingaben im Suchfeld zu suchen, gelangen Anwender mit Sprachsuchen in kürzester Zeit gezielt an relevante Informationen.
- **Vorlesefunktion:** Per Sprachbefehl können Anwender sich gewünschte Kapitel der Bedienungsanleitung laut vorlesen lassen. Allerdings eignen sich bestimmte Kapitel nicht zum Vorlesen, wie Handlungsanweisungen. Die Handlungsschritte sind zum einen oftmals nicht passend zum Vorlesen formuliert und zum anderen werden für die Vermittlung der einzelnen Schritte Grafiken benötigt.[96] Hierfür bieten sich Smart Displays an, wie im Kapitel 7.3.2 erläutert wird.

Digitale Bedienungsanleitungen spielen zudem eine wichtige Rolle, wenn Unternehmen virtuelle Assistenten in Form von Chatbots einsetzen. Als Beispiel ist der Automo-

[95] Vgl. Tekom: https://www.tekom.de/technische-kommunikation-das-fach/wichtige-normen-der-technischen-kommunikation/standard-titel

[96] Vgl. C. Rechtien (Februar 2019): https://technischekommunikation.info/schwerpunktthemen/wie-anleitungen-sprechen-lernen-975/

bilhersteller Daimler anzuführen. Dieser nutzt einen sprachbasierten Assistenten namens „Ask Mercedes“. Das Ziel des Unternehmens liegt darin, mit Hilfe des Assistenten, die Informationen aus der Betriebsanleitung eines Fahrzeugs für Kunden zugänglich zu gestalten.[97] Die App „Ask Mercedes“ stellt eine Kombination aus Augmented Reality und Chatbot dar mit der Nutzung von Künstlicher Intelligenz. Neben einer App, ist „Ask Mercedes“ über Messenger-Dienste und über Sprachassistenten, wie Alexa, verfügbar.

Anwender stellen über die App Fragen in Form von Sprach- oder Texteingaben. Die Fragen beziehen sich auf die Fahrzeugnutzung. Der Assistent ist in der Lage, natürliche Sprache zu verstehen und die Fragen in Echtzeit zu beantworten. Realisierbar ist dies durch Algorithmen, die mit Datenmengen trainiert werden. Für die Beantwortung der Fragen werden Text, Bilder und Videos integriert. Als Grundlage dient die Betriebsanleitung des jeweiligen Fahrzeugs. Wenn der Assistent die Frage nicht versteht oder es mehrere Antwortmöglichkeiten gibt, kann dieser Rückfragen stellen. Bei zu komplexen Fragen wird auf die Betriebsanleitung oder auf den Kundenservice verwiesen.[98]

Darüber hinaus lassen sich mit Augmented Reality über die Smartphone-Kamera Objekte im Cockpit des Fahrzeugs erkennen. Diese werden am Smartphone-Display mit den entsprechenden Funktionen wiedergegeben.[99]

Abbildung 10: Chatbot mit Augmented-Reality-Funktionen über App[100]

[97] Vgl. T. Egberts (25. Mai 2020): https://www.weka.de/produktsicherheit/chatbots-technik-dokumentation/

[98] Vgl. ebd.

[99] Vgl. Mercedes-Benz Consulting (2021): https://www.mercedes-benz-consulting.de/cases/projekt-ask-mercedes-br-wir-schaffen-ein-kundenerlebnis/

Mit dem virtuellen Assistenten wird der Zugang zur Betriebsanleitung vereinfacht und ein personalisiertes Kundenerlebnis ermöglicht.[101]

8.3. Kundenservice

Kunden haben den Anspruch, dass Serviceanliegen einfach, intuitiv und schnell durchgeführt werden.[102] Viele Unternehmen setzen im Kundenservice sprach- und textbasierte Chatbots ein, die einen interaktiven Self-Service ermöglichen und einen ortsunabhängigen und jederzeit verfügbaren Ansprechpartner für Kunden darstellen.[103]

Aus einer repräsentativen Studie des Digitalverbands Bitkom im Jahr 2020 ging hervor, dass jedes vierte Unternehmen in Deutschland Chatbots verwendet. Herr Britze, Bereichsleiter für digitale Geschäftsprozesse von Bitkom, zieht folgenden Rückschluss:

> „Im Zuge der Digitalisierung von Prozessen wird die Nutzung von Chatbots weiter ausgebaut. Neue technische Möglichkeiten entwickeln die automatisierte Kommunikation mit dem Kunden und ermöglichen so mehr Effizienz für das Unternehmen. Solche Automatisierungen erbringen signifikante Kostenvorteile und zusätzliche Service-Verbesserungen.“[104]

Eine technische Möglichkeit stellt der Einsatz von Sprachassistenten in Form von Voice-Apps im Kundenservice dar. Der Autor Herr Hörner hat in seinem Buch „Marketing mit Sprachassistenten“ einen Vergleich gezogen zwischen Sprachassistenten in Form von Voice-Apps und persönlichem Telefonkontakt im Kundenservice:

[100] Vgl. Mercedes-Benz Consulting (2021): https://www.mercedes-benz-consulting.de/cases/projekt-ask-mercedes-br-wir-schaffen-ein-kundenerlebnis/

[101] Vgl. G. Walthart: https://media.daimler.com/marsMediaSite/de/instance/ko/Intelligente-Dialogtechnologie-kombiniert-mit-Augmented-Reality-Ask-Mercedes-Der-virtuelle-Assistent-hilft-sofort-weiter.xhtml?oid=30345702

[102] Vgl. J. Hauser und J. Schleier (2020) S. 209

[103] Vgl. H. Henn (Februar 2018): https://marketing-resultant.de/wp-content/uploads/eBook_Chatbots_FEB2018-1.pdf

[104] N. Britze, N. Paulsen (12. November 2020): Jedes vierte Unternehmen in Deutschland nutzt Chatbots

	Telefonkontakt mit Servicemitarbeiter:	**Sprachassistent in Form einer Voice-App:**
Verfügbarkeit	Begrenzt	Jederzeit
Wartezeit	Abhängig von Anrufaufkommen	Keine Wartezeit
Eignung für routinemäßige Anliegen	Hoch	Hoch
Eignung für komplexe Anliegen	Hoch	Niedrig
Fehlerquote im Dialog	Niedrig	Mittel
Auswertung der Servicevorgänge	Teilautomatisiert	Automatisiert

Tabelle 7: Gegenüberstellung eines Servicemitarbeiters mit einem Sprachassistenten[105]

Ein deutlicher Vorteil von Sprachassistenten im Kundenservice liegt darin, dass Kunden jederzeit ihre Anliegen stellen können und sich nicht nach Zeiten des Kundenservices richten müssen. Die direkte Interaktion findet zwischen Kunden und Sprachassistenten ohne Wartezeiten und Telefonschleifen statt.[106] Dabei kann eine hohe Anzahl an Kunden von Sprachassistenten betreut werden. Dies führt zu einer Entlastung der Servicemitarbeiter und einer Kosteneffizienz seitens der Unternehmen.[107]

Wie in der Tabelle zu erkennen ist, sind Sprachassistenten für standardisierte und wiederholende Dialoge geeignet, wie zum Beispiel für:

- Beratung von Produkten
- Informationen über Kundenkonto
- Durchführung von einfachen Buchungen

[105] Vgl. T. Hörner (2019): https://sprachassistenten.marketing/kundenservice/sprachassistenten-versus-callcenter-agent/

[106] Vgl. T. Hörner (2020) S. 159

[107] Vgl. H. Henn (Februar 2018): https://marketing-resultant.de/wp-content/uploads/eBook_Chatbots_FEB2018-1.pdf

Ungeeignet sind Sprachassistenten bei komplexen Serviceanliegen. Hierfür können Sprachassistenten die Kunden an entsprechenden Servicemitarbeiter weiterleiten.

Hinzu kommt, dass sprachbiometrische Authentifizierungen eine zunehmend wichtige Rolle bei der Erkennung von Kunden spielt. Das Profil von Stimmen ist individuell wie bei Fingerabdrücken. Im Vergleich zur Abfrage von Kundennummern oder Beantwortung von Sicherheitsfragen, erfolgt die Identifikation der Kunden über die Stimmen.[108]

Als Beispiel ist das Unternehmen Nuance Communications anzuführen. Nuance ist ein Anbieter von Sprachverarbeitungslösungen mit Spezialisierung auf Künstliche Intelligenz und Cloud-Computing. Das Unternehmen bietet Authentifizierungsverfahren an, wie die Stimmbiometrie. Die Stimme des Kunden wird in Echtzeit ohne zusätzliche Eingabe mit einem gespeicherten Stimmabdruck verglichen. Dieses Verfahren gilt als eine sichere und anwenderfreundliche Authentifizierungsmethode.[109]

8.4. Zusammenfassung

Angesichts der beschriebenen Anwendungsfälle liegt die Schlussfolgerung nahe, dass Sprachassistenten zwar nicht professionelle Übersetzer oder Mitarbeiter im Kundenservice ersetzen können, aber als Unterstützung dienen, wie z.B. bei einfachen routinemäßigen Kundenanliegen. Bei der Verwendung von Bedienungsanleitungen kann von zeitsparenden Sprachsuchen profitiert werden. Für die multimediale Darstellung eignen sich Smart Displays.

[108] Vgl. P. Kabel (2020) S. 140

[109] Vgl. Nuance Communications (2021): https://www.nuance.com/de-de/omni-channel-customer-engagement/security/identification-and-verification.html#standardpage-mainpar_multicolumn_copy_1290446287

9. Experteninterviews

Dieses Kapitel befasst sich mit dem empirischen Teil der Bachelorarbeit. Als qualitative Forschungsmethode wurden fünf Experteninterviews jeweils in schriftlicher Form durchgeführt. Die Methode der Experteninterviews wurde gewählt, da Experten einen aktuellen Blick bezüglich der Potenziale von Sprachassistenten mit Anwendungsfällen in der Technischen Kommunikation geben können.

Im Gegensatz zu standardisierten Befragungen mit vorformulierten Antwortmöglichkeiten, liegt der Vorteil bei qualitativen Interviews darin, dass die Befragten selbst individuell antworten können.[110] Experteninterviews geben daher Einblick auf Einschätzungen, persönliche Meinungen sowie weitere Anregungen.

9.1. Vorstellung der Experten

Jeder Befragte ist Experte zu jeweiligen Themenbereichen. Dabei können Experten auch mehrere Bereiche abdecken. Laut Bogner definiert sich ein Experte wie folgt:

> „Experten lassen sich als Personen verstehen, die sich – ausgehend von einem spezifischen Praxis- oder Erfahrungswissen, das sich auf einen klar begrenzbaren Problemkreis bezieht – die Möglichkeit geschaffen haben, mit ihren Deutungen das konkrete Handlungsfeld sinnhaft und handlungsleitend für Andere zu strukturieren.“[111]

Beim sogenannten Sampling werden die zu befragenden Experten gezielt ausgewählt bezüglich des zu untersuchenden Forschungsgegenstandes.[112] Das Sampling basiert im Rahmen der Bachelorarbeit auf den Forschungsfragen.

In der anknüpfenden Tabelle werden die Experten vorgestellt mit den jeweiligen Schwerpunkten. In der dritten Spalte wird der Zeitpunkt der Versendung und der schriftlichen Beantwortung der Fragebögen dargestellt.

[110] Vgl. J. Strübing (2018) S. 90

[111] A. Bogner et al. (2014) S. 13

[112] Vgl. ebd. 34 f.

Experten:	Schwerpunkt der Interviews:	Zeitraum der Versendung und Antworten der Fragebögen:
Prof. Dr.-Ing. Andreas Attenberger	▪ Einsatzgebiete in der Technischen Kommunikation ▪ Chancen und Risiken ▪ Datenschutz	Vom 11.01.21 bis zum 20.01.21
Ramona Kraus	▪ Fremdsprachenübersetzung ▪ Spracheingabe in der Technischen Dokumentation	Vom 12.01.21 bis zum 21.02.21
Birgit-Maria Hoppe	▪ Fremdsprachenübersetzung ▪ Spracheingabe in der Technischen Dokumentation	Vom 14.01.21 bis zum 26.01.21
Anonymer Experte 1	▪ Einsatzgebiete in der Technischen Kommunikation ▪ Chancen und Risiken	Vom 14.01.21 bis zum 22.01.21
Anonymer Experte 2	▪ Einsatzgebiete in der Technischen Kommunikation ▪ Chancen und Risiken	Vom 18.01.21 bis zum 22.01.21

Tabelle 8: Experten der Interviews[113]

Im Folgenden wird näher auf den beruflichen Hintergrund und für die Bachelorarbeit relevanten Tätigkeitsfelder der Experten eingegangen.

Herr Prof. Dr.-Ing. Andreas Attenberger:

Der diplomierte Medieninformatiker Herr Prof. Dr.-Ing. Attenberger unterrichtet an der Hochschule München im Studiengang Technische Redaktion und Kommunikation das Modul Kommunikations- und Informationstechnologie. Weitere Themengebiete von Herr Prof. Attenberger sind Computer Networks und Cloud Computing sowie eingebettete Systeme und Datenanalyse. Er arbeitet zudem als Forschungsleiter der digitalen Forensik bei ZITiS (Zentrale Stelle für Informationstechnik im Sicherheitsbereich). Der

[113] Eigene Darstellung

Fragebogen ist auf Anwendungsfelder, Chancen und Risiken von Sprachassistenten sowie auf Datenschutz gerichtet.

Frau Ramona Kraus:

Frau Kraus arbeitet bei der 2W Technische Informations GmbH als Abteilungsleiterin für Technische Dokumentation und Übersetzung. Zudem ist sie als Projektmanagerin in den Bereichen Übersetzung und Lektorat tätig. Das Unternehmen ist Dienstleister für Technische Kommunikation, Grafik und Sprachdienstleistungen für Branchen, wie Maschinenbau, Medizintechnik oder Fahrzeugindustrie. Der Fokus ihres Fragebogens liegt auf der Fremdsprachenübersetzung und Spracheingaben in der Technischen Dokumentation.

Frau Birgit-Maria Hoppe:

Frau Hoppe ist zertifizierte Prozessmanagerin und diplomierte Fachübersetzerin mit interkulturellen Soziologie. Ihr Tätigkeitsbereich im Unternehmen Star Group umfasst Consulting, Prozessdesign sowie Projektmanagement für die Sprachmanagement-Systeme. Dies betreffen Aspekte von Terminologie- und Übersetzungsmanagement im Bereich der multilingualen Technischen Dokumentation. Die Schweizer Unternehmensgruppe Star Group ist spezialisiert auf Technische Dokumentationen und Softwareentwicklung. Der Fragebogen spezialisiert sich, wie bei Frau Kraus, auf die Fremdsprachenübersetzung und Spracheingaben in der Technischen Dokumentation.

Anonymer Experte 1:

Die vierte Person des Interviews möchte gerne anonym bleiben. Die Person ist Autor und Geschäftsführer eines Unternehmens, das Programmierkurse anbietet. Die Tätigkeitsfelder der Person richten sich auf Beratung und Schulung zu Themen digitaler Transformationen. Die Person gibt einen Überblick über Einsatzgebiete und Potenziale von Sprachassistenten in der Technischen Kommunikation.

Anonymer Experte 2:

Die fünfte interviewte Person, die ebenfalls anonym bleiben möchte, arbeitet in einem Medienunternehmen im Bereich Digital & Crossmedia. Die Person ist für die Betreuung des Sprachassistenten Alexa verantwortlich. Dieser wird eingesetzt als zusätzlicher Kanal eingesetzt für die Vermittlung sendungsbegleitender Inhalte. Der Fragebogen richtet sich auf die Potenziale und mögliche Anwendungsfälle von Sprachassistenten.

9.2. Vorgehen der Experteninterviews

Die Kontaktanfrage der Experten fand per E-Mail statt. Sie wurden ausführlich darüber informiert, wie und wofür die Daten weiterverarbeitet und genutzt werden. Zwei Experten äußerten den Wunsch, anonym zu bleiben. Zudem wurde darauf hingewiesen, dass jederzeit Kontakt aufgenommen werden kann, wenn Fragen unklar sind oder mehr Informationen benötigt werden.

Die fünf erstellten Fragebögen wurden als PDF-Dokument versendet. In dem Dokument gab es die Möglichkeit in den hervorgehobenen Textfeldern direkt Antworten zu erfassen. Diese wurden vollständig ausgefüllt und schriftlich zurückgeschickt.

9.3. Aufbau der Fragenbögen

Die Fragen wurden auf das Themengebiet der Experten ausgerichtet. Vereinzelte Fragen wiederholten sich, um einen anderen Blick bestimmter Themen zu erhalten. Zur Orientierung und Struktur wurden die Fragebögen mit einem Leitfaden konzipiert. Angefangen wurde mit einer einleitenden Frage zum generellen Nutzen bzw. aktuellen Stand von Sprachassistenten. Im Hauptteil wurde sich mit den jeweiligen Schwerpunkten der Experten beschäftigt und eine abschließende Schlussfrage im Hinblick auf die Zukunft rundete das schriftliche Interview ab.

Als Fragetypen wurden sich auf erzählungsgenerierende Fragen und Stellungnahmen festgelegt. Anhand von offenen Erzählungen ermöglicht die erste Frageform einen Gewinn von neuen Erkenntnissen des Themengebiets und eine Erschließung von bisher nicht bedachten Aspekten und Sichtweisen der Experten. Stellungnahmen zielen dagegen auf Einschätzungen und Beurteilungen ab.[114]

9.4. Auswertung der Fragebögen

Die Interviews wurden nicht transkribiert, da die Antworten bereits schriftlich vorlagen. Lediglich einzelne Rechtschreibfehler wurden korrigiert. Bei unklaren oder unpräzisen Antworten, wurden telefonisch Rückfragen gestellt, die daraufhin ausführlich und verständlich beantwortet worden sind.

Bei der Auswertung der Interviews erfolgte ein Kodierungsverfahren. Darunter versteht man die Zuordnung von Kategorien bestimmter Textabschnitte. Das Kodieren fand auf Ebene der schriftlichen Einzelinterviews statt. Den Expertenantworten wurden Kodes zu Textabschnitten hinzugefügt, die die Kernaussage beschrieben. Danach erfolgte ein

[114] Vgl. A. Bogner et al. (2014) S. 62 ff.

Vergleich der Kodes aller Interviews, die abschließend übergeordneten Kategorien zugeordnet wurden. Dies ermöglicht, neben einer strukturierten Darstellungsweise, dass gleiche oder ähnliche Antworten zwischen den Experten einfacher und schneller erkennbar sind.[115] Die Fragebögen sowie das angewandte Kodierungsverfahren der Fragebögen befinden sich im Anhang.

9.5. Zusammenfassung der ausgewerteten Experteninterviews

In diesem Kapitel werden die Ergebnisse der Experteninterviews zusammengefasst, die aus dem Kodierungsverfahren hervorgehen. Die Erkenntnisse sind thematisch gegliedert. Die anonymisierten Experten werden als Experte 1 und Experte 2 aufgeführt.

9.5.1. Zielgruppe von Sprachassistenten

Die Experten sind sich einig: Sprachassistenten erreichen eine breitgefächerte Zielgruppe und tragen zur Inklusionsförderung bei. Die Technologie eignet sich für Personen, für die andere Eingabegeräte, wie die Maus oder die Tastatur, eine Barriere darstellen. Besonders ältere Personen, Sehbeeinträchtigte und Personen mit motorischen Einschränkungen können von der Technologie im privaten und beruflichen Alltag profitieren. Allerdings sollte bedacht werden, dass hör- und sprechbeeinträchtigte Personen auf visuelle Schnittstellen gebunden sind.

9.5.2. Stärken und Nutzung von Sprachassistenten

Auch bei diesem Aspekt sind die Experten einer Meinung. Sprachassistenten eignen sich durch die Eingabe natürlicher Sprache für folgende Funktionen:

- Multimediafunktionen (Abspielen von Musik, Hörbücher und Filmen)
- Suchfunktionen (Suche über eine Spracheingabe statt Texteingabe)
- Diktier- und Vorlesefunktion (als Unterstützung von beeinträchtigten Personen)
- Kommunikationsfunktion (Schreiben von Kurznachrichten und Annahme von Telefonanrufen über Sprachbefehle)

Bei einer regelmäßigen Interaktion mit einem Sprachassistenten werden eine steigende Akzeptanz sowie eine persönliche Bindung erzielt. Dies ist vor allem bei personalisierten und erfolgreichen Sprachdialogen der Fall. Zudem können mehrere Anwender

[115] Vgl. A. Bogner et al. (2014) S. 78 f.

einen Sprachassistenten verwenden durch die Hinterlegung von Stimmprofilen verschiedener Anwender. Diese Aspekte steigern die User-Experience.[116]

9.5.3. Schwächen von Sprachassistenten

Zu den Schwächen von Sprachassistenten zählt die Informationseinschränkung. Nur eine bestimmte Anzahl an Informationen können auditiv Anwender vermittelt werden. Bedacht werden muss, dass Sprachassistenten ungeeignet sind, wenn komplexe Parameter eingegeben werden müssen oder die einzugebenden Zeichen schwer auszusprechen sind. Aber auch wenn Umgebungen keine akustisch geeigneten Bedienungen vorweisen oder sensible Daten nicht an Öffentlichkeit sollen. Die noch nicht ausgereifte Technologie der Sprachverarbeitung führt dazu, dass noch einige Missverständnisse bzw. falsche Antworten zwischen Anwender und Sprachassistenten auftreten können.

Die größte Schwäche, bei der alle Experten einer Meinung sind, stellt der Datenschutz dar. Der Sprachassistent ist in der Regel immer empfangsbereit und kann jegliche Daten aufnehmen. Dies sei auch der Grund, weshalb viele Personen sehr kritisch gegenüber Sprachassistenten sind. Anwender sind sich oftmals unsicher, wie und wo die Daten verarbeitet werden und inwiefern das Nutzungsverhalten gesteuert wird.

Herr Prof. Attenberger ergänzt zum Thema Datenschutz, dass aus Medien und technischen Untersuchungen bekannt ist, dass ungewollte Sprachaufzeichnungen in die Cloud geschickt und teilweise von Mitarbeitern abgehört werden. Das sollte bei einer „anonymen“ maschinen-basierten Verarbeitung nicht geschehen und ist für Anwender abschreckend. Ein weiteres Problem stellt die unbewusste Nutzung mit einem Sprachassistenten dar. Anwender wissen oftmals nicht, dass sie aktiv einwilligen können, ob Sprachaufzeichnung gespeichert werden. Herr Prof. Attenberger ergänzt diesbezüglich:

> „Dies ist vielen Kunden aber gar nicht bewusst, manche finden erst spät mit großen Schrecken heraus, wie lückenlos die „Sprachhistorie“ ist. Für mich ist dies übrigens ebenfalls ein Grund, zuhause keinen Sprachassistenten zu nutzen, insbesondere ist nicht gewährleistet, dass nur die expliziten Anfragen übertragen werden. Beispielsweise kann es immer wieder zu Fehlaktivierungen kommen und es besteht natürlich das Potenzial von Hackerangriffen.“[117]

[116] Vgl. Experte 1 (2021) Anhang, S. 80

[117] A. Attenberger (2021) Anhang, S. 69 f.

Hierfür bietet sich als Lösung einerseits eine transparente Aufklärung der Hersteller an und andererseits eine bewusste Nutzung aus Sicht der Anwender (Einwilligung oder Ablehnung der Datenverarbeitung).

9.5.4. Anwendungsmöglichkeiten in der Technischen Kommunikation

Die Experten sind sich einig: Durch die hohen Potenziale von Sprachassistenten erschließen sich zahlreiche Anwendungsmöglichkeiten in der Technischen Kommunikation. Insbesondere für Tätigkeiten bei denen:

- Hände und Augen anderweitig beschäftigt sind
- Keine Tastaturen zur Verfügung stehen
- Erschwerte Umgebungsbedingungen herrschen (z.B. Hitze, Schmutz etc.)
- Handschuhe getragen werden müssen

Sprachassistenten rücken in verschiedenen Berufsfeldern in den Vordergrund, wie z.B. für Mechaniker, die während einer Reparatur über Sprache Schritt für Schritt angeleitet werden oder sich in Echtzeit über Messdaten der Maschine informieren. Zudem können Sprachassistenten bei der Inbetriebnahme oder beim Erlernen von neuen technischen Systemen eingesetzt werden, wie z.B. bei Fertigungs- und Produktionsmaschinen. Des Weiteren könnte man Sprachassistenten nutzen, bei weiteren Fragen, die sich nicht aus den bereitgestellten schriftlichen bzw. audiovisuellen Materialien ergeben. Aber auch bei Handlungsanweisungen sowie zum Protokollieren von Mess- und Einstellwerten und für die Erfassung von Feedback.

Auch Technische Redakteure könnten beispielsweise über einen Sprachbefehl nach benötigten Informationen für die Technische Dokumentation fragen. Hierbei würde der verbale oder schriftliche Austausch zwischen Technischen Redakteuren und Konstrukteuren oder Ingenieuren entfallen. Voraussetzung ist, dass die notwendigen Informationen in einem für den Sprachassistenten zugänglichen Server eingebettet sind.

Die Experten sehen großes Potenzial für Sprachassistenten im Bereich der Echtzeit-Übersetzungen. Mit guter Spracherkennung, könne man sich ohne Sprachkenntnisse in anderen Kulturkreisen zurechtfinden. Speziell im Fahrzeug gestalten sich Assistenzsysteme als sehr nützlich und steigern die Sicherheit des Fahrers, der sich auf keine visuelle Schnittstelle konzentrieren muss.

Einsatz von Sprachassistenten in der Technischen Dokumentation:

Frau Kraus ist der Meinung, dass Sprachassistenten derzeit technisch noch nicht in der Lage sind, Anwender sinnvoll zu unterstützen. Frau Kraus äußert sich wie folgt:

> „Kritisch sehe ich hier, dass Dokumentationen oft schrittweise gelesen werden. Eine Bedienungsanleitung liest man üblicherweise satz- oder absatzweise und führt die Schritte am Gerät gleichzeitig aus. Ein Sprachassistent wird hier (zumindest aktuell) nicht schlau genug sein, die Schritte intelligent abzugrenzen und damit den Anwender optimal zu unterstützen. So könnte der Anwender schnell genervt sein, wenn er immer wieder stoppen muss, wenn es ihm zu schnell geht.“[118]

Allerdings können Sprachassistenten bei Sprachsuchen und als Vorlesefunktion unterstützen. Statt einen Begriff im Stichwortverzeichnis zu suchen, führen Sprachsuchen bei digitalen Anleitungen dazu, dass relevante Informationen schnell und gezielt gefunden werden. Des Weiteren können per Sprachbefehl gewünschte Kapitel laut vorgelesen werden, um so parallel Tätigkeiten auszuführen.

Gemäß Frau Hoppe, eignen sich Sprachassistenten bezüglich der Fremdsprachenübersetzung im Bereich der Technischen Dokumentation nicht. Je nach Produkt und Herkunft des Herstellers müssen Dokumentationen, wie Bedienungsanleitungen, in der Landessprache vorliegen.

Bei der Fremdsprachenübersetzung von Technischer Dokumentationen können haftungsrechtliche Fragen entstehen, da die automatisch generierte Übersetzung eines Sprachassistenten nicht vom Hersteller des jeweiligen Produkts überprüft wird.

Erfolgt ein Unfall des Anwenders aufgrund einer unklaren bzw. einer falschen Übersetzung des Sprachassistenten, können folgende Beteiligte für die Produkthaftung in Frage kommen:

- die Anwender
- der Hersteller des Produkts
- der Hersteller des Sprachassistenten

Dieses Szenario zieht den Rückschluss, dass vor der Nutzung geklärt sein muss, dass sich der Anwender mit der Verwendung einer automatischen Übersetzung durch einen

[118] R. Kraus (2021) Anhang, S. 71

Sprachassistenten aus dem Rahmen der Produkthaftung hinausbewegt und unter Umständen Garantieansprüche (z.B. bei falscher Wartung) verliert.[119]

Einsatz von Sprachassistenten im Kundenservice

Experte 1 ist davon überzeugt, dass Sprachassistenten sinnvoll im Kundenservice eingesetzt werden können. Den Kunden einen zusätzlichen Kanal anzubieten, sieht er als Chance. Der automatisierte Prozess trägt zu einer Entlastung der Servicemitarbeiter bei. Die Künstliche Intelligenz ist in der Lage, Kundenauthentifizierungen sowie routinierte Kundenanfragen auszuführend. Prof. Dr. Attenberger sieht die Kundenauthentifizierung kritisch an, denn dabei ist Frage, wie gut die Erkennungsraten arbeiten. Außerdem könnte der Datenschutz ein Problem aufwerfen.

Experte 2 erkennt keine Chance von Sprachassistenten im Kundenservice. Die Technologie ist noch nicht weit genug ausgereift, um komplexe Anfragen zum einen zu verstehen und zum anderen umzusetzen. Zudem wünschen sich viele Kunden eine persönliche Betreuung im Kundenservice von Mensch zu Mensch. Des Weiteren besteht die Gefahr, dass Sprachassistenten die Anfrage nicht richtig verstehen und möglicherweise falsch ausführen.

9.5.5. Zukunftsblick

Experte 1 schließt sich zwar der Aussage an, dass das nächste Interface für die digitale Welt die Stimme sein wird, aber der entscheidende Durchbruch bisher nicht erfolgte und teilt sich mit Herr Prof. Attenberger die Ansicht, dass der Datenschutz und die Angst vor Überwachung die größten Risiken darstellen. Viele Unternehmen sind nicht transparent genug bezüglich Datenverarbeitung. Laut Herr Prof. Attenberger sei eine zukünftige Entwicklung, anonyme maschinen-basierte Sprachassistenten mit einer optimierten Spracherkennung ohne Aufzeichnung und Weiterverarbeitung von Spracheingaben. Experte 2 zieht folgenden Rückschluss:

> „Wenn sich die Technologie dahingehend weiterentwickelt, dass sich eine darunterliegende KI an den Nutzer und dessen Bedürfnisse anpassen kann, also anfängt, ihn wirklich zu verstehen, sehe ich vor allem im privaten Bereich großes Potenzial. Ein Sprachassistent wäre im besten Fall ein persönlicher Sekretär und ein guter Kollege der den Nutzer in möglichst vielen Situationen das Leben erleichtert. Davon sind wir aber noch ein gutes Stück entfernt.“[120]

[119] Vgl. B. Hoppe (2021) Anhang S. 77

[120] Experte 2 (2021) Anhang, S. 81

9.6. Zusammenfassung

Aus den Erkenntnissen der Experteninterviews lässt sich für die Potenziale und Anwendungsfälle von Sprachassistenten in der Technischen Kommunikation folgendes interpretieren: Alle Experten sind der Meinung, dass Sprachassistenten für den privaten Gebrauch deutliche Vorteile mit sich bringen.

Zudem können unterschiedliche Berufsfelder davon profitieren, wenn es zum Beispiel um die Eingabe und das Protokollieren von Mess- und Einstellwerten geht oder Feedback erfasst werden muss. Für den Einsatz im Kundenservice sind sie gespaltener Meinung. Einerseits ist die automatische Ausführung und Kundenauthentifizierung effizient, andererseits geht dadurch die persönliche Betreuung verloren. Für die Technische Dokumentation sind Sprachassistenten für die Fremdsprachenübersetzung ungeeignet, aufgrund haftungsrechtlicher Aspekte. Der Datenschutz und die Überwachungsgefahr stellen Risiken dar. Diese Risiken können mögliche Gründe sein, warum viele Personen die Nutzung von Sprachassistenten als kritisch sehen. Diesbezüglich sollte eine Aufklärung der Anwender stattfinden.

10. Fazit und Ausblick

Aus Basis der Literaturarbeit, der Produkt- und Funktionsanalyse am Beispiel des Sprachassistenten Alexa und der Experteninterviews können abschließend Rückschlüsse auf die Potenziale von Sprachassistenten mit Anwendungsfällen in der Technischen Kommunikation gezogen werden.

10.1. Fazit

Sprachassistenten bieten bedeutende Chancen für Anwender. Die Technologie erreicht eine breitgefächerte Zielgruppe und stellt eine Unterstützung und Erleichterung des Alltags dar, insbesondere für Personen mit Beeinträchtigungen. Durch die sprachliche Interaktion wird eine einfache, schnelle und zugleich komfortable Bedienung sichergestellt.

Für die Nutzung von Sprachassistenten werden weder Vorwissen noch physische Betätigungen benötigt. Anwender können sich parallel auf andere Tätigkeiten fokussieren und sich frei im Raum bewegen. Sprachassistenten sind jederzeit empfangsbereit und antworten in kürzester Zeit. Diese Chancen bieten viele Einsatzmöglichkeiten von Sprachassistenten in der Technischen Kommunikation, die für Anwender Mehrwerte darstellen. Als Beispiel seien schnelle Sprachsuchen bei langen Dokumentationen, Fernsteuerungen des Fahrzeugs, Fremdsprachenübersetzungen in Echtzeit oder Kundenservice ohne Warteschlagen genannt. Zu berücksichtigen gilt, dass die Intentionen der Anwender im Fokus stehen sollten, um ein positives Nutzererlebnis zu ermöglichen.

Allerdings bringen Sprachassistenten auch Risiken mit, wie z.B. die mangelnde Transparenz hinsichtlich der Datenverarbeitung sowie die Fehleranfälligkeit des Systems. Im Sprachdialog können Missverständnisse und falsche Ausführungen des Sprachbefehls zu einer schnellen Frustration der Anwender führen. Oftmals ist die Nutzung eingeschränkt, wenn Umgebungen keine akustisch geeigneten Bedingungen vorweisen oder sensible Daten an die Öffentlichkeit gelangen sollen. Hinzu kommt, dass nur eine begrenzte Anzahl an Informationen auditiv vermittelt werden kann. Hierfür bietet sich die Erweiterung eines Bildschirms an, wie es bei Smartphones, Tablets oder Smart Displays der Fall ist. Mit dieser Thematik beschäftigt sich Hypothese 3. Im Folgenden werden die zu Beginn aufgestellten Hypothesen bewertet:

Hypothese 1	Natural Language Processing gilt als Schlüsseltechnologie für Sprachassistenten.

Die Technik von Sprachassistenten ist in der Lage, anhand von Erfahrungen zu lernen und personalisierte Dialoge mit Anwendern zu führen. Je häufiger der Sprachassistent in Gebrauch ist, desto besser gestaltet sich das Verständnis. Um eine Interaktion zwischen Anwendern und Sprachassistenten zu ermöglichen und mittels Algorithmen natürliche Sprache zu erkennen, verstehen und darauf zu reagieren, gilt NLP als Schlüsseltechnologie von Sprachassistenten. Dadurch ist Hypothese 1 als verifiziert anzusehen.

Hypothese 2	Anwender gelangen mit der intuitiven Spracheingabe in der Technischen Dokumentation schnell und gezielt an Informationen.

Spracheingaben zählen zu den natürlichsten Interaktionen mit technischen Geräten. Zudem sind Spracheingaben dreimal schneller und 20 Prozent weniger fehleranfällig als Texteingaben. Für die Technische Dokumentation bedeutet dies, dass Anwender mit Hilfe von intuitiven Spracheingaben, wie z.B. Sprachsuchen, in kürzester Zeit und gezielt an relevante Inhalte gelangen. Somit ist Hypothese 2 verifiziert.

Hypothese 3	Bei der Vermittlung von Handlungsanweisungen eignen sich Smart Displays.

Der Funktionsumfang eines Smart Displays ist größer als der eines Smart Speakers durch die Erweiterung eines Bildschirms. Daraus ergeben sich weitere nützliche Funktionen für Anwender, wie die Darstellung multimedialer Inhalte. Die Visualisierung gestaltet sich bei Bedienungsanleitungen als wichtig, um Inhalte effizient zu vermitteln. Bei Handlungsanweisungen bieten sich Videos und Tutorials an, die über Smart Displays abrufbar sind. Anwender können mittels Sprache intuitiv Befehle erteilen und parallel Handlungsschritte durchführen. Dadurch ist Hypothese 3 verifiziert.

Hypothese 4	Mit Sprachassistenten können Mitarbeiter im Kundenservice ersetzt werden.

Sprachassistenten im Kundenservice unterstützen Kunden und entlasten Servicemitarbeiter. Sie können bei routinierten Kundenanfragen zum Einsatz kommen sowie Kunden zum zuständigen Servicemitarbeiter weiterleiten. Außerdem bieten sie eine Unterstützung bei der Kundenauthentifizierung mittels Stimmbiometrie. Allerdings können Sprachassistenten Servicemitarbeiter nicht ersetzen, denn zum einen legen viele Kunden Wert auf eine persönliche Betreuung und zum anderen ist die Technologie nicht in

der Lage, komplexe Anfragen zu verstehen und auszuführen. Deshalb ist Hypothese 4 als falsifiziert anzusehen.

Hypothese 5	Mit Sprachassistenten ergeben sich neue Möglichkeiten in der Fremdsprachenübersetzung.

Sprachassistenten bieten ein hohes Potenzial in der Fremdsprachenübersetzung, die in Echtzeit ausgeführt wird. Neben dem Übersetzen einzelner Wörter ist es möglich, satzübergreifende Zusammenhänge in verschiedenen Sprachen zu übersetzen. Hinzu kommt eine automatische Erkennung der Sprache. Allerdings kann keine fehlerfreie Übersetzung garantiert werden. Einen technologischen Fortschritt bieten maschinelle Übersetzungen mit Verwendung Künstlicher Intelligenz. Es ist eine Frage der Zeit, wann Sprachassistenten davon profitieren können. Daher ist Hypothese 5 differenziert zu betrachten.

Die Bachelorarbeit zeigt, dass, Sprachassistenten, die richtig eingesetzt werden, große Potenziale für Anwender bieten und viele sinnvolle Einsatzmöglichkeiten in der Technischen Kommunikation ermöglichen.

„Der Einsatz von neuen Technologien ist vergleichbar mit dem Einsatz von Werkzeugen. Richtig eingesetzt sollen sie unterstützen, helfen und das Leben erleichtern.“[121]

10.2. Ausblick

Die Stimme wird eine weitere Eingabemöglichkeit in der Mensch-Computer-Interaktion darstellen. Durch die einfache, bequeme und schnelle Bedienung werden Sprachassistenten in vielen Bereichen eine elementare Rolle spielen, wie im Fahrzeug, im Kundenservice oder in der Echtzeit-Übersetzung. Vor allem das Zusammenspiel zwischen Sprachassistent und Bildschirm bietet Potenziale. Der Datenschutz ist bei der Nutzung von Sprachassistenten von hoher Relevanz. Dabei steht die gesellschaftliche Frage im Vordergrund, wie personenbezogene Daten verarbeitet werden. Mit der fortschreitenden Entwicklung der Sprachverarbeitung und Künstlichen Intelligenz werden Sprachassistenten in der Zukunft performanter, intelligenter und weitere Anwendungsfälle bieten.

[121] M. Kaiser et al. (2019) S. 11

Abbildungsverzeichnis

Abbildung 1: Entwicklung der Mensch-Computer-Schnittstellen 6

Abbildung 2: Integrierte CUIs in unterschiedlicher Hardware 9

Abbildung 3: Geräteentwicklung mit integrierten Sprachassistenten 11

Abbildung 4: Nutzung von Sprachassistenten 13

Abbildung 5: Einsatzgebiete von Sprachassistenten 14

Abbildung 6: Dialogablauf zwischen Anwender und Sprachassistenten 20

Abbildung 7: Steigende Intelligenz von Sprachassistenten 25

Abbildung 8: Häufige Fehlerquellen eines VUIs 26

Abbildung 9: Gerätebeschreibung des Smart Displays Echo Show 10 33

Abbildung 10: Chatbot mit Augmented-Reality-Funktionen über App 40

Tabellenverzeichnis

Tabelle 1: Hypothesen und zugehörige Forschungsfragen 2

Tabelle 2: Funktionsweise der Technologie ASR 17

Tabelle 3: Klassifizierung der Technologie NLP 19

Tabelle 4: Gegenüberstellung von Stärken und Schwächen von Sprachassistenten 30

Tabelle 5: Grundlegende Funktionen des Sprachassistenten Alexa 32

Tabelle 6: Technische Details des Smart Displays „Echo Show 10“ 34

Tabelle 7: Gegenüberstellung eines Servicemitarbeiters mit einem Sprachassistenten 42

Tabelle 8: Experten der Interviews 45

Literaturverzeichnis

Albrecht, U. (2020). *Alexa und Amazon Echo.* Berlin: Stiftung Warentest.

Amazon. (2021a). *Mercedes-me*. Abgerufen am 05. März 2021 von Amazon: https://www.amazon.de/Mercedes-Benz-Mercedes-me/dp/B075435C1Y

Amazon. (2021b). *Alexa kennenlernen*. Abgerufen am 03. März 2021 von Amazon: https://www.amazon.de/b?ie=UTF8&node=12775495031

Amazon. (2021c). *Echo Show 10*. Abgerufen am 05. März 2021 von Amazon: https://www.amazon.de/der-neue-echo-show-10-hochauflosendes-smart-display-mit-bewegungsfunktion-und-alexa-anthrazit/dp/B084P3KP2S

Amazon. (2021d). *Häufig gestellte Fragen zu Alexa und Alexa-Geräten*. Abgerufen am 12. März 2021 von Amazon. Hilfe und Kundenservice: https://www.amazon.de/gp/help/customer/display.html?nodeId=201602230#GUID-1CDA0A16-3D5A-47C1-9DD8-FDEDB10381A3__SECTION2

Baumann, M. (11. November 2019). *Amazon Echo Show: Alle Smart Display Modelle im Test-Überblick*. Abgerufen am 15. März 2021 von Home & Smart: https://www.homeandsmart.de/amazon-echo-show-im-test

Bendel, O. (19. Februar 2018). *Chatbot*. Abgerufen am 01. Februar 2021 von Gabler Wirtschaftslexikon: https://wirtschaftslexikon.gabler.de/definition/chatbot-54248/version-277297

Beyto. (März 2020). *Beyto Smart Speaker Studie 2020 Deutschland.* Abgerufen am 23. Februar 2021 von Beyto GmbH: https://mcusercontent.com/cab8c0a7e9a3fef590fd49ce1/files/3cf2966b-5b46-4f79-a597-80fcdb04fcff/Beyto_Smart_Speaker_Studie_2020_Studienreport.pdf

Bogner, A., Littig, B., & Menz, W. (2014). *Interviews mit Experten. Eine praxisorientierte Einführung.* Wiesbaden: Springer Fachmedien.

Brandt, M. (25. April 2018). *Smart Assistenten werden smarter*. Abgerufen am 13. Februar 2021 von Statista:

https://de.statista.com/infografik/9615/allgemeinwissen-von-smart-assistants/

Britze, N., & Paulsen, N. (12. November 2020). *Jedes vierte Unternehmen in Deutschland nutzt Chatbots*. Abgerufen am 15. Februar 2021 von Bitkom: https://www.bitkom.org/Presse/Presseinformation/Jedes-vierte-Unternehmen-in-Deutschland-nutzt-Chatbots

Campbell, I. C. (14. Dezember 2020). *Amazon adds live translation to Alexa's toolbox of skills*. Abgerufen am 05. April 2021 von The Verge: https://www.theverge.com/2020/12/14/22174561/amazon-alexa-live-translation-google-assitant-interpreter-mode

Der Bundesbeauftragte für den Datenschutz und die Informationsfreiheit. (07. 03 2017). *Miniaturisierung der Informations- und Kommunikationstechnik*. Abgerufen am 21. Februar 2021 von bfdi.bund: https://www.bfdi.bund.de/DE/Datenschutz/Themen/Technische_Anwendungen/TechnischeAnwendungenArtikel/Miniaturisierung.html?cms_templateQueryString=sprachassistenten&cms_sortOrder=score+desc

Deutscher Bundestag. (29. Mai 2019). *Zulässigkeit der Transkribierung und Auswertung von Mitschnitten der Sprachsoftware "Alexa" durch Amazon.* Abgerufen am 14. Februar 2021 von Deutscher Bundestag: https://www.bundestag.de/resource/blob/650728/3f72e6abc1c524961e5809002fe20f21/WD-10-032-19-pdf-data.pdf

DIN EN ISO 9241-210:2020-03. (März 2020). Ergonomie der Mensch-System-Interaktion - Teil 210: Menschzentrierte Gestaltung interaktiver Systeme (ISO 9241-210:2019); Deutsche Fassung.

Duden Rechtschreibung. (2021). *Phonem*. Abgerufen am 04. März 2021 von Duden Rechtschreibung: https://www.duden.de/rechtschreibung/Phonem

Egberts, T. (25. Mai 2020). *Chatbots und Technik-Dokumentation*. Abgerufen am 06. April 2021 von Weka Media GmbH: https://www.weka.de/produktsicherheit/chatbots-technik-dokumentation/

Google Assistant-Hilfe. (2021). *Unterhaltungen mit Dolmetschermodus*. Abgerufen am 05. April 2021 von Google Assistant-Hilfe: https://support.google.com/assistant/answer/9234753#zippy=%2Csprachen-f%C3%BCr-den-dolmetschermodus---smartphones-und-

tablets%2Csprachen-f%C3%BCr-den-dolmetschermodus---lautsprecher-smart-displays-und-intelligente-wecker

Gründinger, W., Hofmann, D., & Klett, D. (April 2019). *Sprachassistenten im Smart Home.* (B. D. e.V., Hrsg.) Abgerufen am 02. Februar 2021 von Bundesverband Digitale Wirtschaft (BVDW) e.V.: https://www.bvdw.org/fileadmin/bvdw/upload/publikationen/smart_home/Sprachassistenten_im_SmartHome.pdf

Hellige, H. D. (2008). *Mensch-Computer-Interface. Zur Geschichte und Zukunft der Computerbedienung.* (H. D. Hellige, Hrsg.) Bielefeld: Transcript (Kultur- und Medientheorie).

Henn, H. (Februar 2018). *Chatbots & AI im Customer Service.* Abgerufen am 07. April 2021 von Marketing Resultant GmbH: https://marketing-resultant.de/wp-content/uploads/eBook_Chatbots_FEB2018-1.pdf

Heuer, C. (2021). *Brain-Computer-Interfaces*. Abgerufen am 15. Januar 2021 von Fraunhofer-Institut für Naturwissenschaftlich-Technische Trendanalysen INT: https://www.int.fraunhofer.de/de/geschaeftsfelder/corporate-technology-foresight/Brain-Computer-Interfaces.html

Heuzeroth, T. (21. Februar 2021). *"Das wollen wir ändern" - Warum Amazon sich jetzt von Blinden helfen lässt*. Abgerufen am 03. März 2021 von Welt: https://www.welt.de/wirtschaft/webwelt/article226788181/Mit-Spracherkennung-Amazon-laesst-sich-von-Blinden-helfen.html

Hoffmann, A. (2018). *Chatbots. Einführung in die Zukunft von Marketing, PR und CRM.* Haar: Franzis.

Hörner, T. (2019). *Marketing mit Sprachassistenten. So setzen Sie Alexa, Google Assistant & Co strategisch erfolgreich ein.* Wiesbaden: Springer Fachmedien.

Hörner, T. (2019). *Sprachassistenten versus persönlicher Callcenter-Agenten*. Abgerufen am 07. April 2021 von Marketing mit Sprachassistenten: https://sprachassistenten.marketing/kundenservice/sprachassistenten-versus-callcenter-agent/

Hörner, T. (17. August 2019). *Wie funktionieren Sprachassistenten?* Abgerufen am 05. März 2021 von Marketing mit Sprachassistenten:

https://sprachassistenten.marketing/allgemein/wie-funktionieren-sprachassistenten/

Huber, S., & Litzel, N. (28. Februar 2018). *Was ist ein Chatbot?* Abgerufen am 03. 02 2021 von Bigdata-Insider: https://www.bigdata-insider.de/was-ist-ein-chatbot-a-690591/

Indiegogo. (2021). *Lily: The 1st Smart Speaker that Teaches Chinese*. Abgerufen am 03. April 2021 von Indiegogo: https://www.indiegogo.com/projects/lily-the-1st-smart-speaker-that-teaches-chinese#/

Juhl, D. (2015). *Technische Dokumentation. Praktische Anleitungen und Beispiele* (3. Ausg.). Berlin, Heidelberg: Springer .

Kabel, P. (2020). *Dialog zwischen Mensch und Maschine. Conversational User Interfaces, intelligente Assistenten und Voice-Systeme.* Wiesbaden: Springer Fachmedien.

Kahle, T., & Meißner, D. (2020). *All About Voice. Konzeption, Design und Vermarktung von Anwendungen für digitale Sprachassistenten.* Freiburg, München, Stuttgart: Haufe Group.

Kaiser, M., Hagenauer, J., & Buttkereit, A.-F. (2019). *Journalistische Praxis: Chabots. Automatisierte Kommunikation im Journalismus und in der Public Relation.* Wiesbaden: Springer Fachmedien.

Kinsella, B., & Mutchler, A. (November 2018). *Voice Assistant Consumer Adoption Report.* Abgerufen am 13. Februar 2021 von Voicebot.ai: https://voicebot.ai/wp-content/uploads/2019/01/voice-assistant-consumer-adoption-report-2018-voicebot.pdf

Kreutzer, R., & Seyed Vousoghi, D. (2020). *Voice-Marketing. Der Siegeszug der digitalen Assistenten.* Wiesbaden: Springer Fachmedien.

Krösmann, C., & Klöß, S. (24. Juli 2020). *Beliebte Helfer: Sprachassistenten haben sich durchgesetzt*. Abgerufen am 15. Februar 2021 von Bitkom: https://www.bitkom.org/Presse/Presseinformation/Beliebte-Helfer-Sprachassistenten-haben-sich-durchgesetzt#:~:text=Beliebte%20Helfer%3A%20Sprachassistenten%20haben%20sich%20durchgesetzt,-

2%20von%205&text=Berlin%2C%2024.%20Juli%202020%20%2D,Zuruf%20%20%E2%80%9

Landay, J., Wobbrock, J., & Liou, K. (23. August 2006). *Speech is 3x Faster than Typing for English and Mandarin Text Entry on Mobile Devices.* Abgerufen am 11. Februar 2021 von Research Gate: https://hci.stanford.edu/research/speech/paper/speech_paper.pdf

Leukert, B., Müller, J., & Noga, M. (2019). Künstliche Intelligenz. Mit Algorithmen zum wirtschaftlichen Erfolg. In P. Buxmann, & H. Schmidt (Hrsg.), *Das intelligente Unternehmen: Maschinelles Lernen mit SAP zielgerichtet einsetzen* (S. 41-62). Berlin, Heidelberg: Springer Gabler.

Mercedes-Benz Consulting. (2021). *Ein smarter Assistent*. Abgerufen am 06. April 2021 von Mercedes-Benz Consulting GmbH: https://www.mercedes-benz-consulting.de/cases/projekt-ask-mercedes-br-wir-schaffen-ein-kundenerlebnis/

Monitor Deloitte. (November 2018). Beyond Touch - Voice Commerce 2030. Wie Voice-assisted Interfaces den Handel in Europa revolutionieren werden.

Myers, C. M., Furqan, A., & Nebosky, J. (April 2018). *Pattern for How Users Overcome Obstacles in Voice User Interfaces.* Abgerufen am 14. Februar 2021 von Research Gate: https://www.researchgate.net/publication/324670825_Patterns_for_How_Users_Overcome_Obstacles_in_Voice_User_Interfaces

Nuance Communications. (2021). *Kundenauthentifizierung mit Biometrie*. Abgerufen am 07. April 2021 von Nuance Communications: https://www.nuance.com/de-de/omni-channel-customer-engagement/security/identification-and-verification.html#standardpage-mainpar_multicolumn_copy_1290446287

Olschewski, M. (07. Februar 2019). *Google Übersetzungsfunktion nutzen - so geht's*. Abgerufen am 05. April 2021 von Home & Smart: https://www.homeandsmart.de/google-home-dolmetscherfunktion

Olson, C., & Kemery, K. (2019). *Voice report. From answers to action: customer adoption of voice technology and digital assistants.* Abgerufen am 16. Februar 2021 von Microsoft: https://advertiseonbing-

blob.azureedge.net/blob/bingads/media/insight/whitepapers/2019/04%20apr/voice-report/bingads_2019_voicereport.pdf

Pfister, B., & Kaufmann, T. (2017). *Sprachverarbeitung. Grundlagen und Methoden der Sprachsynthese und Spracherkennung* (2. Ausg.). Wiesbaden: Springer .

Pflege.digital. (2021). *Digitale Assistenten in der Pflege*. Abgerufen am 18. März 2021 von Pflege.digital: https://www.pflege.digital/

Pickell, D. (14. Februar 2019). *What ist a voice assistant and are they the future of chatbots?* Abgerufen am 13. Februar 2021 von Learning Hub: https://learn.g2.com/voice-assistant

Rechtien, C. (Februar 2019). *Wie Anleitungen sprechen lernen.* Abgerufen am 12. April 2021 von Technische Kommunikation: https://technischekommunikation.info/schwerpunktthemen/wie-anleitungen-sprechen-lernen-975/

Retresco. (2021a). *Natural Language Generation: Aus Daten werden Texte*. Abgerufen am 12. Februar 2021 von Retresco: https://www.retresco.de/lexikon/natural-language-generation/

Retresco. (2021b). *Was ist Natural Language Understanding*. Abgerufen am 11. Februar 2021 von Retresco: https://www.retresco.de/lexikon/natural-language-understanding/

Richter, M., & Flückiger, M. (2016). *Usability und UX kompakt. Produkte für Menschen* (4. Ausg.). Berlin, Heidelberg: Springer .

Ryte Wiki. (2020a). *Digitale Sprachassistenten*. Abgerufen am 02. Februar 2021 von Ryte: https://de.ryte.com/wiki/Digitale_Sprachassistenten

Ryte Wiki. (2020b). *Natural Language Processing*. Abgerufen am 04. März 2021 von Ryte: https://de.ryte.com/wiki/Natural_Language_Processing

Ryte Wiki. (2020c). *User Interface*. Abgerufen am 02. Februar 2021 von Ryte: https://de.ryte.com/wiki/User_Interface#Voice_User_Interface_VUI

Ryte Wiki. (2020d). *User-Centered-Design*. Abgerufen am 19. Januar 2021 von Ryte: https://de.ryte.com/wiki/User_Centered_Design

Ryte Wiki. (2020e). *Voice Search*. Abgerufen am 11. März 2021 von Ryte: https://de.ryte.com/wiki/Voice_Search

Schlegel, H. (29. Juni 2020). *Postbank Digitalstudie 2020. Immer mehr Deutsche sprechen mit digitalen Sprachassistenten.* Abgerufen am 03. Februar 2021 von Postbank: https://www.presseportal.de/pm/6586/4637002

Schleier, J., & Hauser, J. (2020). CRM goes digital. Digitale Kundenschnittstellen in Marketing, Vertrieb und Service exzellent gestalten und nutzen. In D. D. Laux, M. Pufahl, & M. Stadelmann (Hrsg.), *Disruptive Kundeninteraktion durch Stimmen- und Spracherkennung für Banken und Versicherungen* (S. 207-218). Wiesbaden: Springer Fachmedien.

Schmalz, A. (2019). Maschinelle Übersetzung. In V. Wittpahl (Hrsg.), *Künstliche Intelligenz* (S. 194-211). Berlin, Heidelberg: Springer.

Schmidt, C. (27. Dezember 2019). *Das WAV-Audioformat*. Abgerufen am 04. März 2021 von Canto: https://www.canto.com/de/blog/wav-audioformat/

Siepermann, M. (19. Februar 2018). *Digital Native*. Abgerufen am 02. Februar 2021 von Gabler Wirtschaftslexikon: https://wirtschaftslexikon.gabler.de/definition/digital-native-54496

Strübing, J. (2018). *Qualitative Sozialforschung. Eine komprimierte Einführung* (2. Ausg.). Berlin, Boston: De Gruyter Oldenbourg.

Tekom Deutschland e.V (a). (kein Datum). *Die Bedeutung von Nutzungsinformationen in elektronischer Form*. Abgerufen am 05. April 2021 von Tekom Deutschland e.V.: https://www.tekom.de/technische-kommunikation-das-fach/wichtige-normen-der-technischen-kommunikation/standard-titel

Tekom Deutschland e.V. (b). (kein Datum). *Intelligente Informationen*. Abgerufen am 10. Januar 2021 von Tekom Deutschland e.V.: https://www.tekom.de/technische-kommunikation-das-fach/intelligente-information

Tekom Deutschland e.V. (c). (kein Datum). *Was ist Technische Kommunikation?* Abgerufen am 10. Januar 2021 von Tekom Deutschland e.V.: https://www.tekom.de/technische-kommunikation-das-fach/was-ist-technische-kommunikation

Walthart, G. (22. November 2017). *Intelligente Dialogtechnologie kombiniert mit Augmented Reality: Ask Mercedes: Der virtuelle Assistent hilft sofort weiter*. Abgerufen am 06. April 2021 von DaimlerAG: https://media.daimler.com/marsMediaSite/de/instance/ko/Intelligente-Dialogtechnologie-kombiniert-mit-Augmented-Reality-Ask-Mercedes-Der-virtuelle-Assistent-hilft-sofort-weiter.xhtml?oid=30345702

Wendel, M. (19. Mai 2019a). *Wie funktioniert Alexa? Bauweise und Funktionen einfach erklärt*. Abgerufen am 03. März 2021 von Home & Smart: https://www.homeandsmart.de/wie-funktioniert-alexa

Wendel, M. (07. Dezember 2020b). *Alexa Datenschutz: Wie viel Sicherheit bietet Amazon Alexa?* Abgerufen am 22. Februar 2021 von Home & Smart: https://www.homeandsmart.de/alexa-datenschutz-oder-datenspion

Wendel, M. (07. Dezember 2020c). *Was ist Zigbee und welche Zigbee Produkte gibt es?* Abgerufen am 15. März 2021 von Home & Smart: https://www.homeandsmart.de/zigbee-funkprotokoll-hausautomation

Wendel, M. (11. Januar 2021d). *Alexa App einrichten und Amazon Echo effektiv nutzen*. Abgerufen am 03. März 2021 von Home & Smart: https://www.homeandsmart.de/how-to-amazon-alexa-app-effektiv-nutzen

Wendel, M. (14. April 2021e). *Echo Show 10 Test - so gut ist das innovative Alexa Display*. Abgerufen am 16. 04 2021 von Home & Smart: https://www.homeandsmart.de/echo-show-10-test

Wulf, D. (18. Dezember 2018). *Amazon spricht über Alexa und Datenschutz*. Abgerufen am 24. Februar 2021 von Home & Smart: https://www.homeandsmart.de/alexa-datenschutz-interview-amazon

Anhang

I. Fragebögen der Experten 68

Fragebogen 1 68

Fragebogen 2 71

Fragebogen 3 74

Fragebogen 4 78

Fragebogen 5 80

II. Auswertung der Fragebögen mit Hilfe des Kodierungsverfahrens 82

I. Fragebögen der Experten

Fragebogen 1

Experteninterview mit Prof. Dr. Andreas Attenberger

Mit der fortschreitenden Entwicklung der Künstlichen Intelligenz haben digitale Sprachassistenten stark an Bedeutung gewonnen. Der Trend geht immer mehr in Richtung Voice User Interfaces – sprachgesteuerte Benutzerschnittstellen. Insbesondere im Bereich der Technischen Kommunikation können Sprachassistenten einen erheblichen Mehrwert für Nutzer als auch für Unternehmen bieten und in vielen Anwendungsfällen zum Einsatz kommen.

Frage 1: Bei der Interaktion zwischen Mensch und Computer rückt zunehmend die natürliche Stimme in den Fokus. Die Technologie dahinter NLP (Natural Language Processing) weist hohes Potenzial auf. Welche Anwendungsfelder profitieren Ihrer Meinung nach von dieser Technologie?

Alle Steuerungsdienste, bei denen man gerne „hands-free“ agiert (also Assistenten-funktionen beispielsweise im Fahrzeug, in der Küche, Wohnzimmer) zur Steuerung insbesondere von Multimediafunktionen (Musik/Video abspielen) wie auch Such-dienste sowie Messaging/Telefonie (Schreiben von Kurznachrichten, Annahme oder Start von Telefonanrufen, etc.). Denkt man an die Pandemie dann sicherlich auch „digitale Wegweiser“ und Assistenten beispielsweise bei öffentlichen Dienstleistern oder Firmen. Außerdem könnte es noch mehr für Diktierfunktionen genutzt werden oder in der Unterstützung von Leuten mit Handicap. Dafür wird aber wahrscheinlich noch weitere Forschung und Verbesserung der Erkennungsqualität notwendig sein.

Frage 2: Wie können Sprachassistenten in der Technischen Kommunikation sinnvoll zum Einsatz kommen (wie z.B. Echtzeit-Übersetzungen oder Voice-Biometrie)?

Echtzeit-Übersetzungen sind ein tolles Thema aber es ist verdammt schwer, „gut“ zu übersetzen. Vielleicht ist es aber irgendwann soweit, dass man sich in anderen Kulturkreisen ohne Sprachkenntnisse einigermaßen zurechtfinden kann. Sprachassistenten wären auch denkbar bei Inbetriebnahme oder Erlernen neuer technischer Systeme (z.B. Fertigungs/Produktionsmaschinen). Außerdem kann man Sprachassistenten nutzen, um z.B. weitere Fragen, die sich nicht aus den bereitgestellten schriftlichen/audiovisuellen Materialien ergeben zu erfassen und z.B. schriftlich an einen Infodienst weiterzuleiten. Voice-Biometrie wäre zwar denkbar, hier ist mir aber nicht klar, wie sicher die Erkennungsraten sind und da gibt es sicherlich ein Datenschutzproblem. Allerdings ließe sich mit guter Sprechererkennung die Nutzerfreundlichkeit von Assistenzsystemen verbessern,

indem verschiedene Nutzer denselben Sprachassistenten nutzen können ohne sich „in die Quere“ zu kommen.

Frage 3: Eine weitere Art von intelligenter Technikkommunikation sind Geräte, in denen Sprachassistenten implementiert sind. Worin liegen bei der Verwendung die Chancen und Risiken für Nutzer?

Die Chancen sind groß. Die Risiken beziehen sich sicherlich vor allem auf den Datenschutz (wer bekommt welche Daten, wie viele werden gesammelt? Wird dadurch mein Nutzerverhalten gesteuert? Werde ich von einem System, das viel über mich weiß, abhängig?) Ein weiteres Risiko besteht, in dem man bestimmte Nutzergruppen (z.B. Gehörlose) ausschließt. Auf der Chancen-Seite eben die Integration von anderen Nutzern, die einfache Benutzbarkeit von Systemen. Die Unterstützung im Alltag mit nützlichen Diensten. Ich denke, die größte Herausforderung wäre wirklich, das Datenschutz/Sammelproblem zu lösen. Könnte man ei-nen Sprachassistenten bauen, der sehr gute Erkennung anbietet ohne dabei alle Aufnahmen ins Internet zu schicken, wäre das eine richtig tolle zukünftige Entwicklung.

Frage 4: Sind Sie der Meinung, dass Suchanfragen sowie die Beschaffung von Informationen in Zukunft weitestgehend über Spracheingaben verlaufen werden?

Ich denke nicht komplett, sicherlich in vielen Alltagssituationen (siehe „hands-free“) und ins-besondere im Freizeitbereich, allerdings ist man ja hier an die Sprachgeschwindigkeit gebunden und es gibt ja auch die Nutzergruppen, die so etwas nicht nutzen können. Als zusätzliche Möglichkeit aber sicherlich eine nützliche Sache. Insbesondere im Fahrzeug habe ich die Assistenzsysteme als sehr nützlich und auch sicherheitssteigernd kennengelernt, da man sich hier dann weniger auf andere (visuelle) Schnittstellen konzentrieren muss.

Frage 5: Sie als Forschungsleiter der digitalen Forensik bei ZITiS (Zentrale Stelle für Informationstechnik im Sicherheitsbereich) beschäftigen sich u.a. mit der IT-Sicherheit.
Einer der bekanntesten künstlichen Gesprächspartner ist Alexa von Amazon. Der Sicherheitsaspekt dieses Systems ist umstritten. Viele Nutzer habe Bedenken, dass Privatgespräche abgehört und an Amazon versendet werden. Wie stehen Sie zur Datensicherheit und Datenverarbeitung mit Sprachassistenten wie Alexa?

Tatsächlich ist das für mich das größte Problem. Wir wissen ja mittlerweile aus den Medien und auch aus technischen Untersuchungen, dass viele der Aufzeichnungen ins Netz geschickt werden, teilweise auch von Menschen geprüft werden. Dies entspreche nicht der gängigen Vorstellung einer „anonymen“ maschinen-basierten Verarbeitung. Weiterhin speichert Amazon (zwar mit Einwilligung der Nutzer) die meisten der Anfragen. Dies ist vielen Kunden aber gar nicht bewusst, manche finden erst spät mit großen Schrecken heraus, wie lückenlos die „Sprachhistorie“ ist. Für mich ist dies übrigens ebenfalls ein Grund, zuhause

keinen Sprachassistenten zu nutzen, insbesondere ist nicht gewährleistet, dass nur die expliziten Anfragen übertragen werden (beispielsweise kann es immer wieder zu Fehlaktivierungen kommen und es besteht natürlich das Potential von Hackerangriffen.
Hier hilft nur eine bewusste Nutzung und Aufklärung der Nutzer (Medien- bzw. Nutzungskompetenz).

Einverständniserklärung zum Experteninterview

Ich erkläre mich hiermit bereit, an dem Experteninterview teilzunehmen. Ich bin damit einverstanden, dass die Antworten des Interviews als Material für wissenschaftliche Zwecke im Rahmen der Bachelorarbeit genutzt werden dürfen.

München, 20.01.2021

Interviewerin: Natascha Grundner
Experte: Prof. Dr. Andreas Attenberger
Betreuerin: Prof. Dr. Gertrud Grünwied

Hochschule München
University of
Applied Sciences

Hochschule München für angewandte Wissenschaften
Fakultät 05 – WiSe 2020 / 2021

Fragebogen 2

Experteninterview mit Frau Ramona Kraus

Mit der fortschreitenden Entwicklung der Künstlichen Intelligenz haben digitale Sprachassistenten stark an Bedeutung gewonnen. Der Trend geht immer mehr in Richtung Voice User Interfaces – sprachgesteuerte Benutzerschnittstellen. Insbesondere im Bereich der Technischen Kommunikation können Sprachassistenten einen erheblichen Mehrwert für Nutzer als auch für Unternehmen bieten und in vielen Anwendungsfällen zum Einsatz kommen.

Frage 1: Ist es Ihrer Meinung nach sinnvoll als Leiterin der Technik-Übersetzung bei 2W Sprachmarkt, Sprachassistenten im Bereich der Technischen Dokumentation für Fremdsprachenübersetzungen einzusetzen?

Das Geschäft eines Übersetzungsdienstleisters findet üblicherweise aus Basis schriftlicher Texte statt. Der Einsatz von Sprachassistenten findet hier also keine Anwendung. Aus Perspektive eines Endnutzers macht ein solcher Einsatz theoretisch schon Sinn. Je nach Produkt und Herkunft des Herstellers muss eine technische Dokumentation dem Produkt beiliegen, meist auch in der jeweiligen Landessprache. Insofern wären hier Sprachassistenten hier für den Endkunden nur für die Funktion des Vorlesens interessant. Aktuell schätze ich den Bedarf dieser Funktion bei den Endkunden eher geringer ein. Wer ist schon „zu faul“, um sich einen Text selbst durchzulesen? Aber: Wie immer bei „neuer“ Technologie, findet diese durch die Nutzung an anderer Stelle (z. B. wie aktuell am häufigsten im Haushalt) schnell Einzug in die Gewohnheit der Nutzer und damit könnte ich mir vorstellen, dass die Nutzer von Sprachassistenten wie Alexa bald auch per Gewohnheit bei solchen Anwendungsfällen gerne auf die Technologie zurückgreifen wollen.

Kritisch sehe ich hier, dass technische Dokumentation oft schrittweise gelesen wird. Eine Bedienungsanleitung liest man üblicherweise satz- oder absatzweise und führt die Schritte am Gerät gleichzeitig aus. Ein Sprachassistent wird hier (zumindest aktuell) nicht schlau genug sein, die Schritte intelligent abzugrenzen und damit den Anwender optimal zu unterstützen. So könnte der Anwender schnell genervt sein, wenn er immer wieder stoppen muss, wenn es ihm zu schnell geht.

Frage 2: Welche Vorteile hat die intuitive Spracheingabe für Nutzer und Nutzerinnen in der Technischen Dokumentation?

- Bei Menschen mit Sehbehinderung
- Durch freieres Arbeiten – man hat eine Hand frei und kann diese z. B. für Werkzeug ö. ä. nutzen

- Bei komplexen Vorgängen, bei denen theoretisch mehrere Menschen beteiligt sind und nicht einer mit Vorlesen beschäftigt ist
- Durch schnelleres Finden von relevanten Inhalten – man muss nicht mehr erst im Stichwortverzeichnis suchen, das erledigt der Assistent für einen.

Frage 3: Mit der Google-Übersetzer-App lässt sich mit Hilfe der Smartphone-Kamera und einer Texterkennungssoftware gedruckter Text in Echtzeit in insgesamt 27 Sprachen umwandeln. Können Sie sich vorstellen, dass Nutzer und Nutzerinnen mit ihren Smartphones Bedienungsanleitungen oder Service-Dokumentationen übersetzen und laut vorlesen lassen? Welche Probleme könnten hierbei auftreten?

Maschinelle Übersetzung ist leider immer noch nicht so ausgereift, dass eine fehlerfreie Übersetzung sichergestellt werden kann. Erst vor Kurzem hatten wir einen Fall, in dem DeepL den technischen Begriff „screw“ (war hier als Handlungsanweisung „Schrauben Sie“ gemeint) sehr „verhängnisvoll“ übersetzt hat… Auch wenn sich Übersetzer über so einen Fall köstlich amüsieren können, zeigt es, dass auch führende Anbieter aktuell noch keine Lösung anbieten, die die Komplexität von Sprache und vor allem *Fachsprache* abbilden können. Und hier sehe ich auch im Kontext Technischer Dokumentation das Problem, dass nur sehr spezialisierte Engines, die auf technische Inhalte trainiert werde, die korrekte Übersetzung von fachlicher Terminologie leisten können. Dazu müsste Google also eine eigens programmierte Lösung speziell für die Übersetzung technischer Inhalte anbieten. Die meisten Engines sind allerdings auf allgemeinsprachliche Texte ausgelegt, um einen möglichst breiten Anwenderkreis abzudecken. Solange es also keine eigene Lösung gibt, die auf technische Texte trainiert wird, schätze ich die Gefahr einer Falschübersetzung als sehr hoch ein. Gerade auch in Kombination mit Texterkennung über OCR, wo es eine weitere Quelle für Fehler gibt. Ein falsch erkanntes Wort und ein falsch übersetztes Wort und schon läuft der Anwender Gefahr einer falschen Bediendung. Und das kann bei technischer Dokumentation im schlimmsten Fall gefährlich für den Anwender werden, gerade wenn es sich um kein Fachpersonal handelt, sondern private Anwender.

Frage 4: Was muss bei Texten in der Technischen Dokumentation beachtet werden, um eine korrekte Übersetzung zu gewährleisten (wie z.B. Terminologiemanagement)?

Das Thema „übersetzungsgerechtes Schreiben“ ist ja ein sehr weites Feld. Es nimmt aber gerade für die hier thematisierten Technologien (Sprachassistenten, Texterkennung und maschinelle Übersetzung) eine immer wichtigere Rolle ein: Je standardisierter ein Text ist, desto einfacher ist er maschinell zu verarbeiten. Aktuell wird deswegen auch eine neue DIN-Norm für übersetzungsgerechtes Schreiben entworfen. Folgende Punkte sind wichtig für übersetzungsgerechtes Schreiben:

- Kurze, parallele und verständliche Satzstrukturen
- Aktive Formulierungen
- Konsistente Terminologie
- Konsistente Formulierungen
- Verbalisierung statt Nominalisierung
- Vermeidung von Satzklammern
- Vermeidung von Füllwörtern und unnötigen Attributen
- Vermeidung von Tautologien

Frage 5: „Eine intelligente Sprachsoftware kann einen Fachübersetzer ersetzen.“ Schließen Sie sich dieser Aussage an?

Ganz klar: Aktuell noch nicht, aber irgendwann bestimmt. Aber davon sind wir noch weit entfernt. Das Problem mit der Sprachvarietät (Allgemeinsprache vs. Fachsprache) kann aktuell durch die neuronale maschinelle Übersetzung noch nicht abgebildet werden. Das bedeutet, dass es noch keine Lösung gibt, die in Form *eines* Anbieters und *einer* Engine für maschinelle Übersetzung alle Texte gleich gut übersetzen kann. Die Technologien werden aber stets weiterentwickelt und wenn man sieht, was sich hier in den letzten drei Jahren getan hat, wäre es naiv zu glauben, dass man in weiteren drei Jahren nicht schon dort sein kann. Erst Mitte Februar konnte der deutsche Anbieter für MÜ-Technologie lengoo 20 Mio. US Dollar an Venture Investmentkapital beschaffen. Das zeigt sehr deutlich, wie viel auf die Weiterentwicklung der Technologie gesetzt wird und wie viel auch branchenfremde Investoren daran glauben, dass die Anbieter damit erfolgreich sein werden.

Einverständniserklärung zum Experteninterview

Ich erkläre mich hiermit bereit, an dem Experteninterview teilzunehmen. Ich bin damit einverstanden, dass die Antworten des Interviews als Material für wissenschaftliche Zwecke im Rahmen der Bachelorarbeit genutzt werden dürfen.

Ramona Kraus

München, 21.02.2021

Interviewerin: Natascha Grundner
Expertin: Ramona Kraus
Betreuerin: Prof. Dr. Gertrud Grünwied

Hochschule München
University of
Applied Sciences

Hochschule München für angewandte Wissenschaften
Fakultät 05 – WiSe 2020 / 2021

Fragebogen 3

Experteninterview mit Frau Birgit-Maria Hoppe

Mit der fortschreitenden Entwicklung der Künstlichen Intelligenz haben digitale Sprachassistenten stark an Bedeutung gewonnen. Der Trend geht immer mehr in Richtung Voice User Interfaces – sprachgesteuerte Benutzerschnittstellen. Insbesondere im Bereich der Technischen Kommunikation können intuitive Spracheingaben einen erheblichen Mehrwert für Nutzer als auch für Unternehmen bieten und in vielen Anwendungsfällen zum Einsatz kommen.

Frage 1: Sie sind Diplom-Übersetzerin und arbeiten mit Sprachmanagement-Systemen. Verwenden Sie hierfür Spracheingaben (z.B. für die Suchfunktion)?

Mein Tätigkeitsbereich ist Consulting, Prozessdesign und Projektmanagement für die Sprachmanagement-Systeme der STAR Group. Diese betreffen verschiedene Aspekte von Terminologie- und Übersetzungsmanagement, v.a. im Bereich multilinguale Technische Dokumentation:

- Hier benötigen die Anwender für Suchfunktionen oder gezielte Übersichten zu konkreten Themen zwar auch String-Suchen (die aber auch z.B. in der Terminologie als „Fuzzy“ oder „Root“ weiter qualifizierbar sind).
- Zusätzlich benötigen sie strukturierte Eingabemöglichkeiten (komplexe Filtermöglichkeiten inkl. Ein-/Ausschluss einer Vielzahl von Parametern).
- Die Anwender dieser Systeme nutzen insgesamt ca. 4-7 verschiedene Oberflächensprachen und im Schnitt zwischen 5 und ~40 Inhaltssprachen.
- Der potenzielle Nutzen, das alles mit Spracheingaben für Such- oder andere Funktionen abzudecken, stünde in keinem Verhältnis zu den Entwicklungs- und resultierenden Systemkosten.

Frage 2: Wann sind Texteingaben Ihrer Meinung nach sinnvoller?

- wenn komplexe Parameter gesetzt werden müssen,
- wenn die einzugebenden Zeichen schwer/unterschiedlich auszusprechen sind
- wenn die Eingabesprachen nicht über die Sprachassistenz abgedeckt werden
- wenn die Erfassungsumgebung keine akustisch geeigneten Bedingungen bietet (Umgebungslärm)

Frage 3: Welche Vorteile sehen Sie in der intuitiven Spracheingabe?
Ich sehe die intuitiven Spracheingabe derzeit vor allem in Bereichen als vorteilhaft, in denen:

- die Zielgruppe keine Möglichkeit der Eingabe hat (ggf. Kinder, Menschen mit temporären/permanenten Einschränkungen durch Alter, Krankheit, Unfallsituation
- situationsbedingt keine Tastatur zur Verfügung steht (auch z.B. in Bereichen mit schwierigen Umgebungsbedingungen, z.B. Schmutz Wasser, Hitze…),
- Hände und Augen anderweitig beschäftigt sind (und aus Sicherheitsgründen fokussiert und konzentriert bleiben sollen, z.B. Autofahren oder sonstige Maschinenbedienung),
- Hände dreckig oder in Handschuhen sind, so dass Taste nur schlecht oder gar nicht bedient werden können
- einfach der Aufwand des Umwegs über eine Tastatur nicht gerechtfertigt ist (Licht an/aus, Frage nach aktueller Wettervorhersage)
- klare eindeutige Angaben (Handlungsanweisungen, Fragen, Eingaben z.B. zur Prozesssteuerung) möglich sind, auch z.B. im Feld (bei Servicearbeiten) für einfache Protokollierung von Mess- und Einstellwerten, sowie für die Erfassung von Feedback (Bsp. „Ölaustritt bei Ventil 17“)
- ein Freiraum für Interpretation und Rückfragen toleriert ist (z.B. keine Dringlichkeit bzw. Zeit für Klärung von Fuzzy-Eingaben, Rückfragen/Vorschlägen des Sprachassistenten zur sukzessiven Klärung des Frage-/Eingabe-Ziels.

Frage 4: Was muss bei Texten in der Technischen Dokumentation beachtet werden, um eine korrekte Übersetzung zu gewährleisten?

1. Prinzipiell gelten alle Anforderungen der DIN EN ISO 17100 (Übersetzungsdienstleistungen - Anforderungen an Übersetzungsdienstleistungen (ISO 17100:2015); Deutsche Fassung EN ISO 17100:2015).
2. Sowohl grundsätzlich als auch bezüglich der aktuellen Entwicklungen in der Technischen Dokumentation halte ich die folgenden Aspekte für besonders relevant für die Übersetzung:
   - Fachlicher und sprachlicher Kontext müssen geklärt sein. Eine Ansprechperson für fachliche Fragen sollte genannt sein.
   - Ausgangstexte müssen formal und sprachlich übersetzungsgerecht sein (saubere Struktur, vollständige Sätze/Übersetzungseinheiten – siehe einschlägige Literatur)
   - Die Texte sollten hinsichtlich ausgangssprachlicher Terminologie korrekt und konsistent sein; für die Übersetzung sollte die Terminologie mitgeliefert werden – soweit vorhanden, auch für

die Zielsprache. Eine Ansprechperson für terminologische Fragen sollte genannt sein.

- Die zu übersetzenden Einheiten müssen mit genügend Kontext versehen sein (wobei sichergestellt sein muss, dass der Kontext sowohl vom CAT-Tool als auch vom Übersetzer berücksichtigt, aber von der Übersetzungseinheit unterschieden werden kann):
    - entweder direkt als umgebender Kontext in den zu übersetzenden Daten
    - als Metadaten zur Übersetzungseinheit (z.B. als Ergebnis der Auswertung eines Knowledge Graphen zu Baugruppe/Komponente, Variante/Modell, Informationstyp)
    - als mitgelieferte Visualisierung
    - für maschinelle Übersetzung mit entsprechenden Metadaten
- Visualisierung des Kontexts sollte synchronisiert zum zu übersetzenden Text ohne Mehraufwand des Übersetzers verfügbar sein (synchronisierte PDF/HTML-Darstellungen.
- Bei Multimedia-Übersetzungen sollte der mediale Kontext verfügbar sein:
    - Synchron mitlaufender Film bei Untertitel- oder Filmtextübersetzung (Times Tamps)
    - Link zu Audiodateien
- Bei Software-Lokalisierung und anderen längenbegrenzten Texten sollte das Sizing sichtbar und, wenn möglich, anpassbar sein.

3. Bei der Erstellung von Texten für die Nutzung durch Sprachassistenzsysteme sind in der Technischen Redaktion auch die beiden folgenden Punkte zu beachten:
    - Maschineninterpretierbare semantische Darstellung technischer Informationen werden typischerweise in einem Knowledge Graphen abgebildet, so dass Antworten des Sprachassistenten nicht alle vorbereitet werden müssen, sondern vom Knowledge Graphen abgeleitet werden können
    - Relevante Basis sind Ontologien und Terminologien, so dass unterschiedliche Begrifflichkeiten auf standardisierte Terme und Bezeichnungen zurückgeführt werden können.
    - Bsp. „Wie oft muss ich die Kette von meinem Motorrad fetten?“:
      „fetten“ ist ein Schmiervorgang und „Wie oft“ bezieht sich auf ein Intervall, so dass die Frage automatisch auf die Standardabfrage „Schmierintervall in Betriebsstunden der Kette des vorliegenden Motorrads“ zurückgeführt werden kann.

Frage 5: Mit der Google-Übersetzer-App lässt sich mit Hilfe der Smartphone-Kamera und einer Texterkennungssoftware gedruckter Text in Echtzeit in insgesamt 27 Sprachen umwandeln. Können Sie sich vorstellen, dass Nutzer und Nutzerinnen mit ihren Smartphones Bedienungsanleitungen oder Service-Dokumentationen übersetzen und laut vorlesen lassen? Welche Probleme könnten hierbei auftreten?

Laut vorlesen bei PC-Anwendungen gehört bereits in den Bereich der barrierefreien Nutzung und könnte sich meiner Meinung nach auch vom Smartphone aus zu einem sinnvollen und gern genutzten Angebot entwickeln. Vorlesen inklusive Online-Übersetzung dagegen wirft vor allem haftungsrechtliche Fragen auf, da die automatisch generierte Übersetzung ja nicht vom Hersteller eines Produkts überprüft wurde. Wenn also aufgrund einer unklaren oder falschen Übersetzung eine Fehlmanipulation oder ein Wartungs-/Reparaturfehler passiert, der einen schweren Unfall zur Folge hat, stellt sich die Frage, wer nun dafür haftet: der Nutzer, der Hersteller der Übersetzungs-Engine (Google), der Hersteller des Produkts und der Anleitung in der Quellsprache.
D.h. es muss ggf. vor der Nutzung geklärt sein, ob/dass sich der Nutzer mit der Verwendung einer automatischen Übersetzung aus dem Rahmen der Produkthaftung hinausbewegt und unter Umständen (Bsp. bei falscher/ungenügender Wartung) Garantieansprüche verliert.

Einverständniserklärung zum Experteninterview

Ich erkläre mich hiermit bereit, an dem Experteninterview teilzunehmen. Ich bin damit einverstanden, dass die Antworten des Interviews als Material für wissenschaftliche Zwecke im Rahmen der Bachelorarbeit genutzt werden dürfen.

Freising, 26.01.2021

Interviewerin: Natascha Grundner
Expertin: Birgit M. Hoppe
Betreuerin: Prof. Dr. Gertrud Grünwied

Hochschule München
University of
Applied Sciences

Hochschule München für angewandte Wissenschaften
Fakultät 05 – WiSe 2020 / 2021

Fragebogen 4

Experteninterview mit anonymen Experten 1

Mit der fortschreitenden Entwicklung der Künstlichen Intelligenz haben digitale Sprachassistenten stark an Bedeutung gewonnen. Der Trend geht immer mehr in Richtung Voice User Interfaces – sprachgesteuerte Benutzerschnittstellen. Insbesondere im Bereich der Technischen Kommunikation können Sprachassistenten einen erheblichen Mehrwert für Nutzer als auch für Unternehmen bieten und in vielen Anwendungsfällen zum Einsatz kommen.

Frage 1: In Ihrem Buch zitieren Sie den Digitalexperten und Journalisten Sascha Lobo mit der Aussage: „Das nächste Interface für die digitale Welt wird die Stimme." Schließen Sie sich dieser Meinung an?

Grundsätzlich schließe ich mich der Meinung weiterhin an. Allerdings ist der große Durchbruch bisher ausgeblieben. Interessant ist, dass in einigen Filmen und Serien VUIs eingesetzt werden, wie für die Steuerung im Smart-Home oder zum Abspielen der Musik. VUIs weiterhin in Filmen und in Serien genutzt werden. Die negativen Punkte ist zum einen der Datenschutz und die Überwachungsgefahr.

Frage 2: Inwieweit wurden Spracheingaben und -ausgaben in der Technischen Kommunikation genutzt? Wenn wir den Blick in die Zukunft richten: Wo werden Sprachassistenten in der Technischen Kommunikation in den kommenden Jahren verstärkt zum Einsatz kommen?

Sinnvoll sind Spracheingaben und -ausgaben bei Menschen, die keine Hände frei haben, um ihre Arbeit erledigen zu können, wie z.B. Techniker während Reparaturen. Zukünftig könnte der Sprachassistent eventuell auch die Rolle der Technischen Redakteure übernehmen, wenn es darum geht, dass die benötigten Informationen für die Technische Dokumentation bei Konstrukteuren bzw. Ingenieuren eingeholt werden müssen. Großes Potenzial sehe ich bei der Übersetzung von Fremdsprachen, die von Sprachassistenten ausgeführt wird. Allerdings ist die Technik noch nicht so weit, dass die Systeme in hoher Qualität Texte übersetzen können. Diesbezüglich bin ich auf die Entwicklung sehr gespannt.

Frage 3: Viele Unternehmen, wie Lufthansa, Opel und Zalando verwenden im Kundenservice Chatbots. Denken Sie, dass Sprachassistenten wie Alexa & Co. im Kundenservice realisierbar sind?

Definitiv! Das Sprachverständnis des Chatbots beschränkt sich nicht nur auf eine Texteingabe. Man kann Chatbots auch für Audioeingaben. So haben manche Unternehmen ihren Chatbot „recyclen" können, der als Textversion nicht gut angekommen ist, weil die Zielgruppe lieber sprechen statt tippen wollte. Sprachassistenten könnten auch eine große Rolle spielen. Hier ist nur die Frage, in welcher Form sie im Unternehmen implementiert werden können.

Frage 4: In der Studie von Microsoft im Jahr 2019 ging hervor, dass 52 Prozent der Befragten großes Bedenken hinsichtlich der Datensicherheit haben und 41 Prozent der Meinung sind, Sprachassistenten hören passiv mit und zeichnen Gespräche auf. Wie stehen Sie zum Datenschutz und zur Datenverarbeitung bei Sprachassistenten?

Leider ist das weiterhin ein kritisches Thema und definitiv auch eine der großen Bremsen, die die Verbreitung der Sprachassistenten verhindert. Viele Nutzer haben große Bedenken bezüglich Datenschutzes und -verarbeitung. Transparenz ist wichtig, um Nutzer aufzuklären, Akzeptanz und Vertrauen aufzubauen.

Frage 5: Implementierte Sprachassistenten im Bereich der intelligenten Technikkommunikation bringen Stärken und Schwächen mit. Welche überwiegen für Sie persönlich und was halten Sie von Smart Displays?

Implementierte Sprachassistenten im Smartphone, wie Siri, oder stationäre Sprachassistenten, wie Alexa, bieten viele nützliche Funktionen und erleichtern den Alltag. Die große Stärke liegt in der Bequemlichkeit und Schnelligkeit. Sprechen statt tippen ist wesentlich schneller und bequemer. Außerdem ist die Bedienung von VUIs leichter als bei GUIs. Bei der sprachlichen Benutzerschnittstelle müssen nur Sprachbefehle erteilt werden. Eine weitere Stärke von Sprachassistenten ist, dass sich bei einem positiven Nutzererlebnis eine persönliche Bindung zum Assistenten entwickelt, besonders wenn der Dialog erfolgreich war und darüber hinaus Mehrwerte entstanden sind. Des Weiteren können Sprachassistenten individuell auf den Nutzer eingehen durch Anlegen von Stimmprofilen. Die Personalisierung sehe ich als sehr wichtig. Neben dem Datenschutz sehe ich die Fehleranfälligkeit als große Schwäche von Sprachassistenten. Falsche Antworten oder Missverständnisse führen zu schlechten Bewertungen oder einer schlechten Nutzererfahrung. Die Informationseinschränkung ist auch eine Schwäche, allerdings trifft das nur bei Smart Speaker zu ohne Display. Ich sehe großes Potenzial in Sprachassistenten, die einen Display kombinieren, wie Smart Displays. Bei Geräten ohne Display wissen viele Nutzer nicht, welche Funktionen es gibt. Mit einer zusätzlichen visuellen Schnittstelle werden diese Funktionen grafisch angezeigt. Zusätzlich bieten Bedienelemente eine unterstützende Hilfe.

Interviewerin: Natascha Grundner
Experte: Anonym 1
Betreuerin: Prof. Dr. Gertrud Grünwied

Hochschule München
University of
Applied Sciences

Hochschule München für angewandte Wissenschaften
Fakultät 05 – WiSe 2020 / 2021

Fragebogen 5

Experteninterview mit anonymen Experten 2

Mit der fortschreitenden Entwicklung der Künstlichen Intelligenz haben digitale Sprachassistenten stark an Bedeutung gewonnen. Der Trend geht immer mehr in Richtung Voice User Interfaces – sprachgesteuerte Benutzerschnittstellen. Insbesondere im Bereich der Technischen Kommunikation können Sprachassistenten einen erheblichen Mehrwert für Nutzer als auch für Unternehmen bieten und in vielen Anwendungsfällen zum Einsatz kommen.

Frage 1: Welche Zielgruppe ist für Sprachassistenten geeignet bzw. ungeeignet?

Ich denke, dass sich Sprachassistenten sehr gut für die Menschen eignen, für die andere Eingabegeräte – beispielsweise Maus, Tastatur und Co. – eine Barriere darstellen. Vor allem ältere Menschen werden in Zukunft sicher von dieser Technologie profitieren. Auch in verschiedenen Berufsfeldern werden Sprachassistenten in Zukunft vermutlich eine wichtige Rolle spielen. Beispielsweise könnte eine künstliche Intelligenz, die via Sprache gesteuert wird, einen Chirurgen während eine Operation in Echtzeit mit wichtigen Patientendaten versorgen oder einen Mechaniker während einer Reparatur mit einer Schritt für Schritt Anleitung assistieren. Für Kinder halte ich diese Technologie zum aktuellen Zeitpunkt noch ungeeignet. Häufig haben Sprachassistenten noch Probleme, Befehle richtig zu verstehen oder sie Antworten auf eine Weise, die für Kinder nicht zufriedenstellend ist.

Frage 2: Welche Chancen und Risiken sind mit Sprachassistenten verbunden? Wie ist Ihre Erfahrung mit dem Sprachassistenten Alexa?

Chancen sehe ich für Sprachassistenten in allen Situationen, in denen sie uns Menschen das Leben erleichtern können. Sie könnten uns dabei helfen, den beruflichen Alltag stressfreier oder unser Privatleben angenehmer und unterhaltsamer zu gestalten. Aktuell gibt es aber meiner Meinung nach, bedingt durch technologische Limitierungen, noch sehr wenige, wirklich sinnvolle Anwendungsbereiche.

Risiken sehe ich vor allem im Bereich Datenschutz. Damit ein Sprachassistent funktioniert, ist er in der Regel immer empfangsbereit und kann Daten aufnehmen. Der Nutzer weiß aber in der Regel nicht, welche Daten genau aufgenommen werden, wo sie landen und wie diese verarbeitet werden. Hinzu kommt, dass diese Geräte auch gehackt und dazu eingesetzt werden können, Nutzer auszuspionieren.

Bisher habe ich Alexa nur beruflich genutzt. Ziel war es, die Nutzer mit sendungsbegleitenden Inhalten zu versorgen, die sie via Sprachkommandos abrufen

konnten. Das funktionierte und war eine gute Möglichkeit, um feststellen zu können, ob diese Technologie bei der Zielgruppe zum Einsatz kommt. Privat nutze ich keine Sprachassistenten aufgrund der Überwachungsgefahr.

Frage 3: Wünschen Sie sich bei der Abfrage von Informationen eine zusätzlich visuelle Schnittstelle, wie zum Beispiel Smartphones oder Smart Displays?

Aktuell ja. Ich könnte mir aber vorstellen, dass sich das mit der Zeit ändert, wenn sich die Technologie weiterentwickelt und man sich noch mehr an den Einsatz von Sprachassistenten gewöhnt hat.

Frage 4: Was halten Sie von Sprachassistenten in der Fremdsprachenübersetzung und im Kundenservice?

Sprachassistenten in der Fremdsprachenübersetzung finde ich sehr sinnvoll. Meiner Meinung nach ist das bisher einer der nützlichsten Anwendungsfälle. Ich finde, dass Sprachassistenten im Kundenservice weniger sinnvoll sind. Die persönliche Betreuung im Kundenservice, also von Mensch zu Mensch, ist äußerst wichtig, denn der Kunde möchte ernst genommen und für jedes seiner Probleme eine Lösung haben. Das kann ein Sprachassistent aktuell noch nicht bieten, wodurch schnell ein Frusterlebnis entsteht. Die Technik müsste für das Verstehen und für die Umsetzung der Kundenanfrage weiterentwickelt werden.

Frage 5: In welchen Anwendungsfällen sehen Sie großes Potenzial für Sprachsteuerung?

Wenn sich die Technologie dahingehend weiterentwickelt, dass sich eine darunterliegende KI an den Nutzer und dessen Bedürfnisse anpassen kann, also anfängt, ihn wirklich zu verstehen, sehe ich vor allem im privaten Bereich großes Potenzial. Ein Sprachassistent wäre im besten Fall ein persönlicher Sekretär und ein guter Kollege, der den Nutzer in möglichst vielen Situationen das Leben erleichtert. Davon sind wir aber noch ein gutes Stück entfernt.

II. Auswertung der Fragebögen mit Hilfe des Kodierungsverfahrens

Kodierung	Kategorie
▪ Ältere Personen ▪ Kinder ▪ Personen mit Sehbehinderungen ▪ Personen mit motorischen Beeinträchtigungen	**Zielgruppe**
▪ Einfachheit ▪ Schnelligkeit ▪ Bequemlichkeit ▪ Unterhaltung ▪ Hände- und blickfrei ▪ Erleichterung im Alltag ▪ Integration in verschiedenen Geräten ▪ Mehrfachnutzung anderer Anwender ▪ Entwicklung einer persönlichen Bindung	**Stärken und Nutzen**
▪ Datenschutz und –verarbeitung ▪ Gefahr von Hacking und Überwachung ▪ Ausschließen von Gehörlosen ▪ Keine Nutzung bei lauten Hintergrundgeräuschen möglich ▪ Keine grafische Oberfläche ▪ Gefahr von Missverständnissen	**Schwächen**
▪ Sprachsuchen und Vorlesefunktion ▪ Verschiedene Berufsfelder ▪ Fremdsprachenübersetzung ▪ Kundenservice ▪ Steuerung im Smart-Home ▪ Steuerung in Fahrzeugen	**Anwendungsfälle in der Technischen Kommunikation**
▪ Bewusste Nutzung ▪ Aufklärung der Anwender ▪ Multimodalität ▪ Fokus auf Bedürfnisse der Anwender ▪ Fokus auf Transparenz und Datenschutz ▪ Fokus auf Technologie	**Zukunftsblick**

Bei der Interaktion zwischen Mensch und Computer rückt zunehmend die natürliche Stimme in den Fokus. Die Technologie dahinter NLP (Natural Language Processing) weist hohes Potenzial auf. Welche Anwendungsfelder profitieren Ihrer Meinung nach von dieser Technologie?

Kodierung:

- **„Hands-free-Situationen“**

Alle Steuerungsdienste, bei denen man gerne „hands-free“ agiert (also Assistenten-funktionen beispielsweise im Fahrzeug, in der Küche, Wohnzimmer) zur Steuerung insbesondere von Multimediafunktionen (Musik/Video abspielen) wie auch Suchdienste sowie Messaging/Telefonie (Schreiben von Kurznachrichten, Annahme oder Start von Telefonanrufen, etc.). Denkt man an die Pandemie dann sicherlich auch „digitale Wegweiser“ und Assistenten beispielsweise bei öffentlichen Dienstleistern oder Firmen. Außerdem könnte es noch mehr für Diktierfunktionen genutzt werden oder in der Unterstützung von Leuten mit Handicap. Dafür wird aber wahrscheinlich noch weitere Forschung und Verbesserung der Erkennungsqualität notwendig sein.

Wie können Sprachassistenten in der Technischen Kommunikation sinnvoll zum Einsatz kommen (wie z.B. Echtzeit-Übersetzungen oder Voice-Biometrie)?

Kodierungen:

- **Echtzeit-Übersetzungen**
- **Inbetriebnahme oder Erlernen neuer Systeme**
- **Voice-Biometrie**

Echtzeit-Übersetzungen sind ein tolles Thema aber es ist verdammt schwer, „gut“ zu übersetzen. Vielleicht ist es aber irgendwann soweit, dass man sich in anderen Kulturkreisen ohne Sprachkenntnisse einigermaßen zurechtfinden kann. Sprachassistenten wären auch denkbar bei Inbetriebnahme oder Erlernen neuer technischer Systeme (z.B. Fertigungs/Produktionsmaschinen). Außerdem kann man Sprachassistenten nutzen, um z.B. weitere Fragen, die sich nicht aus den bereitgestellten schriftlichen/audiovisuellen Materialien ergeben zu erfassen und z.B. schriftlich an einen Infodienst weiterzuleiten. Voice-Biometrie wäre zwar denkbar, hier ist mir aber nicht klar, wie sicher die Erkennungsraten sind und da gibt es sicherlich ein Datenschutzproblem. Allerdings ließe sich mit guter Sprechererkennung die Nutzerfreundlichkeit von Assistenzsystemen verbessern, indem verschiedene Nutzer denselben Sprachassistenten nutzen können ohne sich „in die Quere“ zu kommen.

Eine weitere Art von intelligenter Technikkommunikation sind Geräte, in denen Sprachassistenten implementiert sind. Worin liegen bei der Verwendung die Chancen und Risiken für Nutzer?

Kodierungen:

- **Datenschutz**
- **Ausschließen von Gehörlosen**
- **Integration von anderen Nutzern**
- **Einfache Benutzbarkeit**
- **Erleichterung im Alltag**
- **Sprachassistent ohne Datenverarbeitung**

Die Chancen sind groß. Die Risiken beziehen sich sicherlich vor allem auf den Datenschutz (wer bekommt welche Daten, wie viele werden gesammelt? Wird dadurch mein Nutzerverhalten gesteuert? Werde ich von einem System, das viel über mich weiß, abhängig?) Ein weiteres Risiko besteht, in dem man bestimmte Nutzergruppen (z.B. Gehörlose) ausschließt. Auf der Chancen-Seite eben die Integration von anderen Nutzern, die einfache Benutzbarkeit von Systemen. Die Unterstützung im Alltag mit nützlichen Diensten. Ich denke, die größte Herausforderung wäre wirklich, das Datenschutz/Sammelproblem zu lösen. Könnte man einen Sprachassistenten bauen, der sehr gute Erkennung anbietet ohne dabei alle Aufnahmen ins Internet zu schicken, wäre das eine richtig tolle zukünftige Entwicklung.

Sind Sie der Meinung, dass Suchanfragen sowie die Beschaffung von Informationen in Zukunft weitestgehend über Spracheingaben verlaufen werden?

Kodierungen:

- **„Hands-free-Situationen“**
- **Sprachsuchen**
- **Assistenz im Fahrzeug**

Ich denke nicht komplett, sicherlich in vielen Alltagssituationen (siehe „hands-free“) und ins-besondere im Freizeitbereich, allerdings ist man ja hier an die Sprachgeschwindigkeit gebunden und es gibt ja auch die Nutzergruppen, die so etwas nicht nutzen können. Als zusätzliche Möglichkeit aber sicherlich eine nützliche Sache. Insbesondere im Fahrzeug habe ich die Assistenzsysteme als sehr nützlich und auch sicherheitssteigernd kennengelernt, da man sich hier dann weniger auf andere (visuelle) Schnittstellen konzentrieren muss.

Sie als Forschungsleiter der digitalen Forensik bei ZITiS (Zentrale Stelle für Informationstechnik im Sicherheitsbereich) beschäftigen sich u.a. mit der IT-Sicherheit. Einer der bekanntesten künstlichen Gesprächspartner ist Alexa von Amazon. Der Sicherheitsaspekt dieses Systems ist umstritten. Viele Nutzer habe Bedenken, dass Privatgespräche abgehört und an Amazon versendet werden. Wie stehen Sie zur Datensicherheit und Datenverarbeitung mit Sprachassistenten wie Alexa?

Kodierungen:

- **Datenschutz**
- **Überwachung**
- **Bewusste Nutzung**
- **Aufklärung der Anwender**

Tatsächlich ist das für mich das größte Problem. Wir wissen ja mittlerweile aus den Medien und auch aus technischen Untersuchungen, dass viele der Aufzeichnungen ins Netz geschickt werden, teilweise auch von Menschen geprüft werden. Dies entspreche nicht der gängigen Vorstellung einer „anonymen" maschinen-basierten Verarbeitung. Weiterhin speichert Amazon (zwar mit Einwilligung der Nutzer) die meisten der Anfragen. Dies ist vielen Kunden aber gar nicht bewusst, manche finden erst spät mit großen Schrecken heraus, wie lückenlos die „Sprachhistorie" ist. Für mich ist dies übrigens ebenfalls ein Grund, zuhause keinen Sprachassistenten zu nutzen, insbesondere ist nicht gewährleistet, dass nur die expliziten Anfragen übertragen werden (beispielsweise kann es immer wieder zu Fehlaktivierungen kommen und es besteht natürlich das Potential von Hackerangriffen. Hier hilft nur eine bewusste Nutzung und Aufklärung der Nutzer (Medien- bzw. Nutzungskompetenz).

Ist es Ihrer Meinung nach sinnvoll als Leiterin der Technik-Übersetzung bei 2W Sprachmarkt, Sprachassistenten im Bereich der Technischen Dokumentation für Fremdsprachenübersetzungen einzusetzen?

Kodierungen:

- **Vorlesefunktion**
- **Fehlende Intelligenz**

Das Geschäft eines Übersetzungsdienstleisters findet üblicherweise aus Basis schriftlicher Texte statt. Der Einsatz von Sprachassistenten findet hier also keine Anwendung. Aus Perspektive eines Endnutzers macht ein solcher Einsatz theoretisch schon Sinn. Je nach Produkt und Herkunft des Herstellers muss eine technische Dokumentation dem Produkt beiliegen, meist auch in der jeweiligen Landessprache. Insofern wären hier Sprachassistenten hier für den Endkunden nur für die Funktion des Vorlesens interessant. Aktuell schätze ich den Bedarf dieser Funktion bei den Endkunden eher geringer ein. Wer ist schon „zu faul", um sich einen Text selbst durchzulesen? Aber: Wie immer bei „neuer" Technologie,

findet diese durch die Nutzung an anderer Stelle (z. B. wie aktuell am häufigsten im Haushalt) schnell Einzug in die Gewohnheit der Nutzer und damit könnte ich mir vorstellen, dass die Nutzer von Sprachassistenten wie Alexa bald auch per Gewohnheit bei solchen Anwendungsfällen gerne auf die Technologie zurückgreifen wollen.

Kritisch sehe ich hier, dass technische Dokumentation oft schrittweise gelesen wird. Eine Bedienungsanleitung liest man üblicherweise satz- oder absatzweise und führt die Schritte am Gerät gleichzeitig aus. Ein Sprachassistent wird hier (zumindest aktuell) nicht schlau genug sein, die Schritte intelligent abzugrenzen und damit den Anwender optimal zu unterstützen. So könnte der Anwender schnell genervt sein, wenn er immer wieder stoppen muss, wenn es ihm zu schnell geht.

Welche Vorteile hat die intuitive Spracheingabe für Nutzer und Nutzerinnen in der Technischen Dokumentation?

Kodierungen:

- **Menschen mit Sehbehinderung**
- **Gewährt freies Arbeiten**
- **Sprachsuche**

  - Bei Menschen mit Sehbehinderung
  - Durch freieres Arbeiten – man hat eine Hand frei und kann diese z. B. für Werkzeug ö. ä. nutzen
  - Bei komplexen Vorgängen, bei denen theoretisch mehrere Menschen beteiligt sind und nicht einer mit Vorlesen beschäftigt ist
  - Durch schnelleres Finden von relevanten Inhalten – man muss nicht mehr erst im Stichwortverzeichnis suchen, das erledigt der Assistent für einen.

Mit der Google-Übersetzer-App lässt sich mit Hilfe der Smartphone-Kamera und einer Texterkennungssoftware gedruckter Text in Echtzeit in insgesamt 27 Sprachen umwandeln. Können Sie sich vorstellen, dass Nutzer und Nutzerinnen mit ihren Smartphones Bedienungsanleitungen oder Service-Dokumentationen übersetzen und laut vorlesen lassen? Welche Probleme könnten hierbei auftreten?

Kodierung:

- **Gefahr von Falschübersetzungen**

Maschinelle Übersetzung ist leider immer noch nicht so ausgereift, dass eine fehlerfreie Übersetzung sichergestellt werden kann. Erst vor Kurzem hatten wir einen Fall, in dem DeepL den technischen Begriff „screw“ (war hier als Hand-

lungsanweisung „Schrauben Sie“ gemeint) sehr „verhängnisvoll“ übersetzt hat. Auch wenn sich Übersetzer über so einen Fall köstlich amüsieren können, zeigt es, dass auch führende Anbieter aktuell noch keine Lösung anbieten, die die Komplexität von Sprache und vor allem *Fachsprache* abbilden können. Und hier sehe ich auch im Kontext Technischer Dokumentation das Problem, dass nur sehr spezialisierte Engines, die auf technische Inhalte trainiert werde, die korrekte Übersetzung von fachlicher Terminologie leisten können. Dazu müsste Google also eine eigens programmierte Lösung speziell für die Übersetzung technischer Inhalte anbieten.

Die meisten Engines sind allerdings auf allgemeinsprachliche Texte ausgelegt, um einen möglichst breiten Anwenderkreis abzudecken. Solange es also keine eigene Lösung gibt, die auf technische Texte trainiert wird, schätze ich die Gefahr einer Falschübersetzung als sehr hoch ein. Gerade auch in Kombination mit Texterkennung über OCR, wo es eine weitere Quelle für Fehler gibt. Ein falsch erkanntes Wort und ein falsch übersetztes Wort und schon läuft der Anwender Gefahr einer falschen Bediendung. Und das kann bei technischer Dokumentation im schlimmsten Fall gefährlich für den Anwender werden, gerade wenn es sich um kein Fachpersonal handelt, sondern private Anwender.

Was muss bei Texten in der Technischen Dokumentation beachtet werden, um eine korrekte Übersetzung zu gewährleisten (wie z.B. Terminologiemanagement)?

Kodierung:

- **Übersetzung in Technischer Dokumentation**

Das Thema „übersetzungsgerechtes Schreiben“ ist ja ein sehr weites Feld. Es nimmt aber gerade für die hier thematisierten Technologien (Sprachassistenten, Texterkennung und maschinelle Übersetzung) eine immer wichtigere Rolle ein: Je standardisierter ein Text ist, desto einfacher ist er maschinell zu verarbeiten. Aktuell wird deswegen auch eine neue DIN-Norm für übersetzungsgerechtes Schreiben entworfen. Folgende Punkte sind wichtig für übersetzungsgerechtes Schreiben:

- Kurze, parallele und verständliche Satzstrukturen
- Aktive Formulierungen
- Konsistente Terminologie
- Konsistente Formulierungen
- Verbalisierung statt Nominalisierung
- Vermeidung von Satzklammern
- Vermeidung von Füllwörtern und unnötigen Attributen
- Vermeidung von Tautologien

„Eine intelligente Sprachsoftware kann einen Fachübersetzer ersetzen.“ Schließen Sie sich dieser Aussage an?

Kodierung:

- **Fremdsprachenübersetzung**

Ganz klar: Aktuell noch nicht, aber irgendwann bestimmt. Aber davon sind wir noch weit entfernt. Das Problem mit der Sprachvarietät (Allgemeinsprache vs. Fachsprache) kann aktuell durch die neuronale maschinelle Übersetzung noch nicht abgebildet werden. Das bedeutet, dass es noch keine Lösung gibt, die in Form *eines* Anbieters und *einer* Engine für maschinelle Übersetzung alle Texte gleich gut übersetzen kann. Die Technologien werden aber stets weiterentwickelt und wenn man sieht, was sich hier in den letzten drei Jahren getan hat, wäre es naiv zu glauben, dass man in weiteren drei Jahren nicht schon dort sein kann. Erst Mitte Februar konnte der deutsche Anbieter für MÜ-Technologie lengoo 20 Mio. US Dollar an Venture Investmentkapital beschaffen. Das zeigt sehr deutlich, wie viel auf die Weiterentwicklung der Technologie gesetzt wird und wie viel auch branchenfremde Investoren daran glauben, dass die Anbieter damit erfolgreich sein werden.

Sie sind Diplom-Übersetzerin und arbeiten mit Sprachmanagement-Systemen. Verwenden Sie hierfür Spracheingaben (z.B. für die Suchfunktion)?

Kodierungen:

- **Sprachsuchen**
- **Keine grafische Oberfläche**

Mein Tätigkeitsbereich ist Consulting, Prozessdesign und Projektmanagement für die Sprachmanagement-Systeme der STAR Group. Diese betreffen verschiedene Aspekte von Terminologie- und Übersetzungsmanagement, v.a. im Bereich multilinguale Technische Dokumentation:

- Hier benötigen die Anwender für Suchfunktionen oder gezielte Übersichten zu konkreten Themen zwar auch String-Suchen (die aber auch z.B. in der Terminologie als „Fuzzy“ oder „Root“ weiter qualifizierbar sind).
- Zusätzlich benötigen sie strukturierte Eingabemöglichkeiten (komplexe Filtermöglichkeiten inkl. Ein-/Ausschluss einer Vielzahl von Parametern).
- Die Anwender dieser Systeme nutzen insgesamt ca. 4-7 verschiedene Oberflächensprachen und im Schnitt zwischen 5 und ~40 Inhaltssprachen.

- Der potenzielle Nutzen, das alles mit Spracheingaben für Such- oder andere Funktionen abzudecken, stünde in keinem Verhältnis zu den Entwicklungs- und resultierenden Systemkosten.

Wann sind Texteingaben Ihrer Meinung nach sinnvoller?

Kodierungen:

- **Umgebungslärm**
- **Komplexe, schwer aussprechbare Eingaben**

- wenn komplexe Parameter gesetzt werden müssen,
- wenn die einzugebenden Zeichen schwer/unterschiedlich auszusprechen sind
- wenn die Eingabesprachen nicht über die Sprachassistenz abgedeckt werden
- wenn die Erfassungsumgebung keine akustisch geeigneten Bedingungen bietet (Umgebungslärm)

Welche Vorteile sehen Sie in der intuitiven Spracheingabe?

Kodierungen:

- **Kinder**
- **Menschen mit Einschränkungen**
- **Autofahren**
- **Maschinenbedienung**
- **Verschmutzte Hände**
- **Handlungsanweisungen**
- **Servicearbeiten**

Ich sehe die intuitiven Spracheingabe derzeit vor allem in Bereichen als vorteilhaft, in denen:

- die Zielgruppe keine Möglichkeit der Eingabe hat (ggf. Kinder, Menschen mit temporären/permanenten Einschränkungen durch Alter, Krankheit, Unfallsituation
- situationsbedingt keine Tastatur zur Verfügung steht (auch z.B. in Bereichen mit schwierigen Umgebungsbedingungen, z.B. Schmutz Wasser, Hitze…),
- Hände und Augen anderweitig beschäftigt sind (und aus Sicherheitsgründen fokussiert und konzentriert bleiben sollen, z.B. Autofahren oder sonstige Maschinenbedienung),
- Hände dreckig oder in Handschuhen sind, so dass Taste nur schlecht oder gar nicht bedient werden können

- einfach der Aufwand des Umwegs über eine Tastatur nicht gerechtfertigt ist (Licht an/aus, Frage nach aktueller Wettervorhersage)
- klare eindeutige Angaben (Handlungsanweisungen, Fragen, Eingaben z.B. zur Prozesssteuerung) möglich sind, auch z.B. im Feld (bei Servicearbeiten) für einfache Protokollierung von Mess- und Einstellwerten, sowie für die Erfassung von Feedback (Bsp. „Ölaustritt bei Ventil 17“)
- ein Freiraum für Interpretation und Rückfragen toleriert ist (z.B. keine Dringlichkeit bzw. Zeit für Klärung von Fuzzy-Eingaben, Rückfragen/Vorschlägen des Sprachassistenten zur sukzessiven Klärung des Frage-/Eingabe-Ziels.

Was muss bei Texten in der Technischen Dokumentation beachtet werden, um eine korrekte Übersetzung zu gewährleisten?

Kodierungen:

- **Übersetzungen in der Technischen Dokumentation**
- **Visualisierung durch GUIs**

1. Prinzipiell gelten alle Anforderungen der DIN EN ISO 17100 (Übersetzungsdienstleistungen - Anforderungen an Übersetzungsdienstleistungen (ISO 17100:2015); Deutsche Fassung EN ISO 17100:2015).
2. Sowohl grundsätzlich als auch bezüglich der aktuellen Entwicklungen in der Technischen Dokumentation halte ich die folgenden Aspekte für besonders relevant für die Übersetzung:
   - Fachlicher und sprachlicher Kontext müssen geklärt sein. Eine Ansprechperson für fachliche Fragen sollte genannt sein.
   - Ausgangstexte müssen formal und sprachlich übersetzungsgerecht sein (saubere Struktur, vollständige Sätze/Übersetzungseinheiten – siehe einschlägige Literatur)
   - Die Texte sollten hinsichtlich ausgangssprachlicher Terminologie korrekt und konsistent sein; für die Übersetzung sollte die Terminologie mitgeliefert werden – soweit vorhanden, auch für die Zielsprache. Eine Ansprechperson für terminologische Fragen sollte genannt sein.
   - Die zu übersetzenden Einheiten müssen mit genügend Kontext versehen sein (wobei sichergestellt sein muss, dass der Kontext sowohl vom CAT-Tool als auch vom Übersetzer berücksichtigt, aber von der Übersetzungseinheit unterschieden werden kann):
     - entweder direkt als umgebender Kontext in den zu übersetzenden Daten
     - als Metadaten zur Übersetzungseinheit (z.B. als Ergebnis der Auswertung eines Knowledge Graphen zu

Baugruppe/Komponente, Variante/Modell, Informationstyp)
    - als mitgelieferte Visualisierung
    - für maschinelle Übersetzung mit entsprechenden Metadaten
- Visualisierung des Kontexts sollte synchronisiert zum zu übersetzenden Text ohne Mehraufwand des Übersetzers verfügbar sein (synchronisierte PDF/HTML-Darstellungen.
- Bei Multimedia-Übersetzungen sollte der mediale Kontext verfügbar sein:
    - Synchron mitlaufender Film bei Untertitel- oder Filmtextübersetzung (Times Tamps)
    - Link zu Audiodateien
- Bei Software-Lokalisierung und anderen längenbegrenzten Texten sollte das Sizing sichtbar und, wenn möglich, anpassbar sein.

3. Bei der Erstellung von Texten für die Nutzung durch Sprachassistenzsysteme sind in der Technischen Redaktion auch die beiden folgenden Punkte zu beachten:
    - Maschineninterpretierbare semantische Darstellung technischer Informationen werden typischerweise in einem Knowledge Graphen abgebildet, so dass Antworten des Sprachassistenten nicht alle vorbereitet werden müssen, sondern vom Knowledge Graphen abgeleitet werden können
    - Relevante Basis sind Ontologien und Terminologien, so dass unterschiedliche Begrifflichkeiten auf standardisierte Terme und Bezeichnungen zurückgeführt werden können.
    - Bsp. „Wie oft muss ich die Kette von meinem Motorrad fetten?“:
      „fetten“ ist ein Schmiervorgang und „Wie oft“ bezieht sich auf ein Intervall, so dass die Frage automatisch auf die Standardabfrage „Schmierintervall in Betriebsstunden der Kette des vorliegenden Motorrads“ zurückgeführt werden kann.

Mit der Google-Übersetzer-App lässt sich mit Hilfe der Smartphone-Kamera und einer Texterkennungssoftware gedruckter Text in Echtzeit in insgesamt 27 Sprachen umwandeln. Können Sie sich vorstellen, dass Nutzer und Nutzerinnen mit ihren Smartphones Bedienungsanleitungen oder Service-Dokumentationen übersetzen und laut vorlesen lassen? Welche Probleme könnten hierbei auftreten?

Kodierungen:

- **Barrierefreie Nutzung**
- **Haftungsrechtliche Probleme**

Laut vorlesen bei PC-Anwendungen gehört bereits in den Bereich der barrierefreien Nutzung und könnte sich meiner Meinung nach auch vom Smartphone aus zu einem sinnvollen und gern genutzten Angebot entwickeln. Vorlesen inklusive Online-Übersetzung dagegen wirft vor allem haftungsrechtliche Fragen auf, da die automatisch generierte Übersetzung ja nicht vom Hersteller eines Produkts überprüft wurde. Wenn also aufgrund einer unklaren oder falschen Übersetzung eine Fehlmanipulation oder ein Wartungs-/Reparaturfehler passiert, der einen schweren Unfall zur Folge hat, stellt sich die Frage, wer nun dafür haftet: der Nutzer, der Hersteller der Übersetzungs-Engine (Google), der Hersteller des Produkts und der Anleitung in der Quellsprache. D.h. es muss ggf. vor der Nutzung geklärt sein, ob/dass sich der Nutzer mit der Verwendung einer automatischen Übersetzung aus dem Rahmen der Produkthaftung hinausbewegt und unter Umständen (Bsp. bei falscher/ungenügender Wartung) Garantieansprüche verliert.

In Ihrem Buch zitieren Sie den Digitalexperten und Journalisten Sascha Lobo mit der Aussage: „Das nächste Interface für die digitale Welt wird die Stimme.“ Schließen Sie sich dieser Meinung an?

Kodierungen:

- **Steuerung im Smart Home**
- **Datenschutz**
- **Überwachungsgefahr**

Grundsätzlich schließe ich mich der Meinung weiterhin an. Allerdings ist der große Durchbruch bisher ausgeblieben. Interessant ist, dass in einigen Filmen und Serien VUIs eingesetzt werden, wie für die Steuerung im Smart-Home oder zum Abspielen der Musik. VUIs weiterhin in Filmen und in Serien genutzt werden. Die negativen Punkte ist zum einen der Datenschutz und die Überwachungsgefahr.

Inwieweit wurden Spracheingaben und -ausgaben in der Technischen Kommunikation genutzt? Wenn wir den Blick in die Zukunft richten: Wo werden Sprachassistenten in der Technischen Kommunikation in den kommenden Jahren verstärkt zum Einsatz kommen?

Kodierungen:

- **„Hands-free-Situationen“**
- **Fremdsprachenübersetzung**

Sinnvoll sind Spracheingaben und -ausgaben bei Menschen, die keine Hände frei haben, um ihre Arbeit erledigen zu können, wie z.B. Techniker während Reparaturen. Zukünftig könnte der Sprachassistent eventuell auch die Rolle der Technischen Redakteure übernehmen, wenn es darum geht, dass die benötigten

Informationen für die Technische Dokumentation bei Konstrukteuren bzw. Ingenieuren eingeholt werden müssen. Großes Potenzial sehe ich bei der Übersetzung von Fremdsprachen, die von Sprachassistenten ausgeführt wird. Allerdings ist die Technik noch nicht so weit, dass die Systeme in hoher Qualität Texte übersetzen können. Diesbezüglich bin ich auf die Entwicklung sehr gespannt.

Viele Unternehmen, wie Lufthansa, Opel und Zalando verwenden im Kundenservice Chatbots. Denken Sie, dass Sprachassistenten wie Alexa & Co. im Kundenservice realisierbar sind?

Kodierung:

- **Kundenservice**

Definitiv! Das Sprachverständnis des Chatbots beschränkt sich nicht nur auf eine Texteingabe. Man kann Chatbots auch für Audioeingaben. So haben manche Unternehmen ihren Chatbot „recyclen" können, der als Textversion nicht gut angekommen ist, weil die Zielgruppe lieber sprechen statt tippen wollte. Sprachassistenten könnten auch eine große Rolle spielen. Hier ist nur die Frage, in welcher Form sie im Unternehmen implementiert werden können.

In der Studie von Microsoft im Jahr 2019 ging hervor, dass 52 Prozent der Befragten großes Bedenken hinsichtlich der Datensicherheit haben und 41 Prozent der Meinung sind, Sprachassistenten hören passiv mit und zeichnen Gespräche auf. Wie stehen Sie zum Datenschutz und zur Datenverarbeitung bei Sprachassistenten?

Kodierungen:

- **Datenschutz**
- **Aufklärung der Nutzer**

Leider ist das weiterhin ein kritisches Thema und definitiv auch eine der großen Bremsen, die die Verbreitung der Sprachassistenten verhindert. Viele Nutzer haben große Bedenken bezüglich Datenschutzes und -verarbeitung. Transparenz ist wichtig, um Nutzer aufzuklären, Akzeptanz und Vertrauen aufzubauen.

Implementierte Sprachassistenten im Bereich der intelligenten Technikkommunikation bringen Stärken und Schwächen mit. Welche überwiegen für Sie persönlich und was halten Sie von Smart Displays?

Kodierungen:

- **Erleichterung im Alltag**
- **Bequemlichkeit**
- **Schnelligkeit**
- **Leichte Bedienung**

- **Entwicklung einer persönlichen Bindung**
- **Personalisierung**
- **Fehleranfälligkeit des Systems**
- **Smart Displays**

Implementierte Sprachassistenten im Smartphone, wie Siri, oder stationäre Sprachassistenten, wie Alexa, bieten viele nützliche Funktionen und erleichtern den Alltag. Die große Stärke liegt in der Bequemlichkeit und Schnelligkeit. Sprechen statt tippen ist wesentlich schneller und bequemer. Außerdem ist die Bedienung von VUIs leichter als bei GUIs. Bei der sprachlichen Benutzerschnittstelle müssen nur Sprachbefehle erteilt werden. Eine weitere Stärke von Sprachassistenten ist, dass sich bei einem positiven Nutzererlebnis eine persönliche Bindung zum Assistenten entwickelt, besonders wenn der Dialog erfolgreich war und darüber hinaus Mehrwerte entstanden sind.

Des Weiteren können Sprachassistenten individuell auf den Nutzer eingehen durch Anlegen von Stimmprofilen. Die Personalisierung sehe ich als sehr wichtig. Neben dem Datenschutz sehe ich die Fehleranfälligkeit als große Schwäche von Sprachassistenten. Falsche Antworten oder Missverständnisse führen zu schlechten Bewertungen oder einer schlechten Nutzererfahrung. Die Informationseinschränkung ist auch eine Schwäche, allerdings trifft das nur bei Smart Speaker zu ohne Display. Ich sehe großes Potenzial in Sprachassistenten, die einen Display kombinieren, wie Smart Displays. Bei Geräten ohne Display wissen viele Nutzer nicht, welche Funktionen es gibt. Mit einer zusätzlichen visuellen Schnittstelle werden diese Funktionen grafisch angezeigt. Zusätzlich bieten Bedienelemente eine unterstützende Hilfe.

Welche Zielgruppe ist für Sprachassistenten geeignet bzw. ungeeignet?

Kodierungen:

- **Menschen mit Beeinträchtigungen**
- **Ältere Menschen**
- **Verschiedene Berufsfelder, wie z.B. Mechaniker**

Ich denke, dass sich Sprachassistenten sehr gut für die Menschen eignen, für die andere Eingabegeräte – beispielsweise Maus, Tastatur und Co. – eine Barriere darstellen. Vor allem ältere Menschen werden in Zukunft sicher von dieser Technologie profitieren. Auch in verschiedenen Berufsfeldern werden Sprachassistenten in Zukunft vermutlich eine wichtige Rolle spielen. Beispielsweise könnte eine künstliche Intelligenz, die via Sprache gesteuert wird, einen Chirurgen während eine Operation in Echtzeit mit wichtigen Patientendaten versorgen oder einen Mechaniker während einer Reparatur mit einer Schritt für Schritt Anleitung assistieren. Für Kinder halte ich diese Technologie zum aktuellen Zeitpunkt noch ungeeignet. Häufig haben Sprachassistenten noch Probleme, Befehle richtig zu

verstehen oder sie Antworten auf eine Weise, die für Kinder nicht zufriedenstellend ist.

Welche Chancen und Risiken sind mit Sprachassistenten verbunden? Wie ist Ihre Erfahrung mit dem Sprachassistenten Alexa?

Kodierungen:

- **Erleichterung im Alltag**
- **Unterhaltung**
- **Hacking**
- **Überwachungsgefahr**

Chancen sehe ich für Sprachassistenten in allen Situationen, in denen sie uns Menschen das Leben erleichtern können. Sie könnten uns dabei helfen, den beruflichen Alltag stressfreier oder unser Privatleben angenehmer und unterhaltsamer zu gestalten. Aktuell gibt es aber meiner Meinung nach, bedingt durch technologische Limitierungen, noch sehr wenige, wirklich sinnvolle Anwendungsbereiche.

Risiken sehe ich vor allem im Bereich Datenschutz. Damit ein Sprachassistent funktioniert, ist er in der Regel immer empfangsbereit und kann Daten aufnehmen. Der Nutzer weiß aber in der Regel nicht, welche Daten genau aufgenommen werden, wo sie landen und wie diese verarbeitet werden. Hinzu kommt, dass diese Geräte auch gehackt und dazu eingesetzt werden können, Nutzer auszuspionieren.

Bisher habe ich Alexa nur beruflich genutzt. Ziel war es, die Nutzer mit sendungsbegleitenden Inhalten zu versorgen, die sie via Sprachkommandos abrufen konnten. Das funktionierte und war eine gute Möglichkeit, um feststellen zu können, ob diese Technologie bei der Zielgruppe zum Einsatz kommt. Privat nutze ich keine Sprachassistenten aufgrund der Überwachungsgefahr.

Wünschen Sie sich bei der Abfrage von Informationen eine zusätzlich visuelle Schnittstelle, wie zum Beispiel Smartphones oder Smart Displays?

Kodierung:

- **Multimodalität**

Aktuell ja. Ich könnte mir aber vorstellen, dass sich das mit der Zeit ändert, wenn sich die Technologie weiterentwickelt und man sich noch mehr an den Einsatz von Sprachassistenten gewöhnt hat.

Was halten Sie von Sprachassistenten in der Fremdsprachenübersetzung und im Kundenservice?

Kodierungen:

- **Fremdsprachenübersetzung**
- **Kundenservice**

Sprachassistenten in der Fremdsprachenübersetzung finde ich sehr sinnvoll. Meiner Meinung nach ist das bisher einer der nützlichsten Anwendungsfälle. Ich finde, dass Sprachassistenten im Kundenservice weniger sinnvoll sind. Die persönliche Betreuung im Kundenservice, also von Mensch zu Mensch, ist äußerst wichtig, denn der Kunde möchte ernst genommen und für jedes seiner Probleme eine Lösung haben. Das kann ein Sprachassistent aktuell noch nicht bieten, wodurch schnell ein Frusterlebnis entsteht. Die Technik müsste für das Verstehen und für die Umsetzung der Kundenanfrage weiterentwickelt werden.

In welchen Anwendungsfällen sehen Sie großes Potenzial für Sprachsteuerung?

Kodierungen:

- **Weiterentwicklung der Technologie**
- **Anpassung an Bedürfnisse der Anwender**

Wenn sich die Technologie dahingehend weiterentwickelt, dass sich eine darunterliegende KI an den Nutzer und dessen Bedürfnisse anpassen kann, also anfängt, ihn wirklich zu verstehen, sehe ich vor allem im privaten Bereich großes Potenzial. Ein Sprachassistent wäre im besten Fall ein persönlicher Sekretär und ein guter Kollege, der den Nutzer in möglichst vielen Situationen das Leben erleichtert. Davon sind wir aber noch ein gutes Stück entfernt.