

1 Verkörperte KI in Form androider Roboter

1.1 Assistenzsysteme und Verkörperung

Moderne, textbasierter Assistenzsysteme haben durch die Nutzung datengetriebener Dienste der künstlichen Intelligenz (KI), wie z.B. *ChatGPT*, ein ganz neues Level erreicht. Allerdings ist die Interaktion auch mit *ChatGPT* auf reine Spracheingabe und -ausgabe, per Text oder Audio, beschränkt. Dieser Einschränkung entgegenzuwirken war bereits vor mehr als zwanzig Jahren das Ziel von Forschenden auf dem Gebiet der sogenannte *Embodied Conversational Agents*. Diese verkörperten Agenten wurden von KI-Forschenden als neuartige Schnittstelle für die Interaktion mit Computern vorgeschlagen (Cassell 2001). Durch die Verkörperung der Schnittstelle wird es dem Nutzer möglich, eine persönlichere Beziehung zum Informationssystem aufzubauen. Dabei sind die verschiedenen Eigenschaften des Designs maßgeblich für die Qualität dieser Beziehung, wie zum Beispiel die äußerliche Gestaltung, die Gestik und Mimik oder die Emotionalität des Agenten (Loveys u. a. 2020).

Parallel zur Entwicklung dieser virtuellen Agenten, die auf dem Computerbildschirm dreidimensional dargestellt und möglichst natürlich animiert werden, wurden in den Neunzigerjahren sogenannte *soziale Roboter* entwickelt. Zuerst wurden Roboter wie COG (Brooks u. a. 1999) oder WABIAN (Hashimoto u. a. 2002; Hashimoto und Takanishi 2019) dem menschlichen Körperbau nachempfunden, da argumentiert wurde, dass unsere Intelligenzleistung zu einem wesentlichen Teil von unserer körperlichen Verankerung in der Umwelt profitiere. Diese Grundidee entstammt dem *Embodiment* (Dautenhahn 1999) genannten Forschungsansatz. Dieser Ansatz stand im Gegensatz zu Systemen, die mittels Logik komplexe Schlüsse über abstrahierte Gegebenheiten der (sozialen) Umwelt schließen, sogenannte Expertensysteme wie z.B. SHAKEY (Fikes und Nilsson 1971; Nilsson 2009; Speck, Dornhege, und Burgard

2017), nur schlecht mit kontinuierlich über Sensorik eintreffenden Daten integriert werden können.

Mit humanoiden Robotern, die dem menschlichen Körperbau physisch und sensorisch nachempfunden sind, besteht seit Ende des letzten Jahrhunderts die Hoffnung, die Integration von kognitiv-abstrakten und reaktiv-körperlichen Intelligenzleistungen entscheidend voranzubringen. Zugleich ergeben sich aber aufgrund der Komplexität dieser Systeme eine Vielzahl neuer Herausforderungen auf den Gebieten der Regelungstechnik und der Robotik. Wenn ein Kleinkind zum Beispiel von allein lernt, mit einem Löffel ein Jogurt zu essen, so ist das aus Sicht der Robotik eine beachtliche, bis heute nicht erreichte körperlich-planerische, sensomotorische Intelligenzleistung. Im Laufe seines Lebens wird das Baby aber sehr wahrscheinlich niemals einen modernen Schachroboter schlagen können (Becker-Asano u. a. 2014).

Eine weitere, wichtige Dimension menschlicher Intelligenz ist die emotional-soziale Intelligenz (Goleman 1995; Lindblom und Ziemke 2003). Das bereits erwähnte Kleinkind entwickelt in den ersten drei Lebensjahren eine interne Repräsentation seiner selbst. Erst dann ist es dazu in der Lage zu verstehen, dass sein Gegenüber eine eigene Perspektive auf die und Interpretation der Umgebung besitzt und Ereignisse dementsprechend subjektiv bewertet. Diese Bewertung von Aktionen und Ereignissen kann zu Emotionen führen, die die Handlungsplanung unter Umständen entscheidend beeinflussen (Becker-Asano 2008; Ortony, Clore, und Collins 1988). Ein Kleinkind kann schon sehr früh über Gesichtsausdrücke und Laute emotional interagieren und Bezugspersonen sind intuitiv in der Lage, diese sozialen Signale zu dekodieren und ihre Handlungen entsprechend anzupassen, die emotionale, non-verbale Rückmeldungen beinhalten (Holodynski und Friedlmeier 2005).

Das eingangs erwähnte Assistenzsysteme *ChatGPT* ist somit in vielfältiger Weise unvollkommen, um als ähnlich mächtiger Gesprächspartner zu fungieren. Ihm fehlt eine Körperlichkeit, die es ihm ermöglichen würde, Gestik und Mimik effizient zur Kommunikation einzusetzen. Es besitzt

außerdem kein Situationsbewusstsein, das es befähigen könnte, an einer dynamisch geführten Diskussion teilzunehmen. Um dieses zu erlangen, müssten Assistenzsysteme über Sensorik verfügen, mit denen sie auch die Gestik und Mimik der Gesprächspartner wahrnehmen können. Diese Wahrnehmung muss in Echtzeit vom System im Kontext der sprachlichen Information ausgewertet werden und schließlich mit den multimodalen Ausdrucksfähigkeiten in Einklang gebracht werden, die dem System selbst ggf. zur Verfügung stehen. Erst ein solches Assistenzsystem könnte einen menschlichen Assistenten vollumfänglich ersetzen. Dabei ist es allerdings umstritten, ob dieses System tatsächlich eine explizite Repräsentation nicht-rationaler Aspekte wie z.B. der Emotionen haben muss, oder nicht (Becker, Kopp, und Wachsmuth 2007; Jung, Müller, und Boll 2022; Mainzer 2019b).

Es liegt nun nahe sich hierbei an Science Fiction Filme wie zum Beispiel „A.I.“ (Spielberg 2001) oder auch „I, Robot“ (Sietz 2004) erinnert zu fühlen. In beiden exemplarisch ausgewählten Filmen wird mit den Urängsten der Zuschauenden gespielt, indem potenzielle Gefahren einer *starken KI* (vgl. (Russell und Norvig 2021), Kapitel 28) durchgespielt werden. Aus technischer Perspektive gibt es derzeit wenig Anlass, Angst vor dieser hypothetischen *Superintelligenz* (Mainzer 2019a) haben zu müssen – egal ob diese in Form einer ubiquitären KI oder eines humanoiden Roboters daher kommt.

1.2 Der Androide Roboter „Andrea“

Androide Roboter sind eine spezielle Art menschenähnlicher Roboter, die einem Menschen zum Verwechseln ähnlichsehen (Ishiguro 2005). Im Gegensatz dazu sind humanoide Roboter zwar ebenfalls dem Menschen („human“) ähnlich („-oid“) gestaltet, aber sie bleiben klar als Roboter erkennbar (Hashimoto und Takanishi 2019). Die Entwicklung androider Roboter begann im Prinzip schon im 18. Jahrhundert mit dem Schachtürken (Standage 2002) durch Wolfgang von Kempelen, einer mechanischen Konstruktion, in der ein Mensch versteckt die Schachzüge eines „mechanischen Türken“ durchführte. Die androiden Roboter der *Geminoid*

Reihe—seit Anfang der 2000er Jahre in Japan entwickelt—wurden zumeist ebenfalls ferngesteuert. Dies geschah allerdings aus der Ferne, was als Tele-Operation bezeichnet wird (Becker-Asano u. a. 2010; Nishio, Ishiguro, und Hagita 2007). Zudem sind Geminoiden Androide immer Kopien real existierender Personen. Erst einige Jahre später wurde der ERICA Roboter als Kunstfigur entworfen und war als einer der ersten androiden Roboter in der Lage, sich autonom mit Menschen auf Japanisch zu unterhalten (Glas u. a. 2016).

Der Idee des Roboters ERICA folgend, wurde der androide Roboter Andrea als autonomer, androgyn gestalteter Roboter an der Hochschule der Medien erschaffen. Auch Andrea ist nicht als Kopie eines realen Menschen gestaltet, sondern das Äußere des Roboters wurde mit einer 3D Modellierungssoftware explizit so entworfen, dass ein Betrachter ihm kein eindeutiges Geschlecht zuordnen kann.



Der Androide Roboter Andrea im Labor des Instituts für angewandte, künstliche Intelligenz der Hochschule der Medien, Stuttgart

Andrea verfügt über insgesamt 52 pneumatische Aktuatoren, die in der folgenden Tabelle aufgelistet sind.

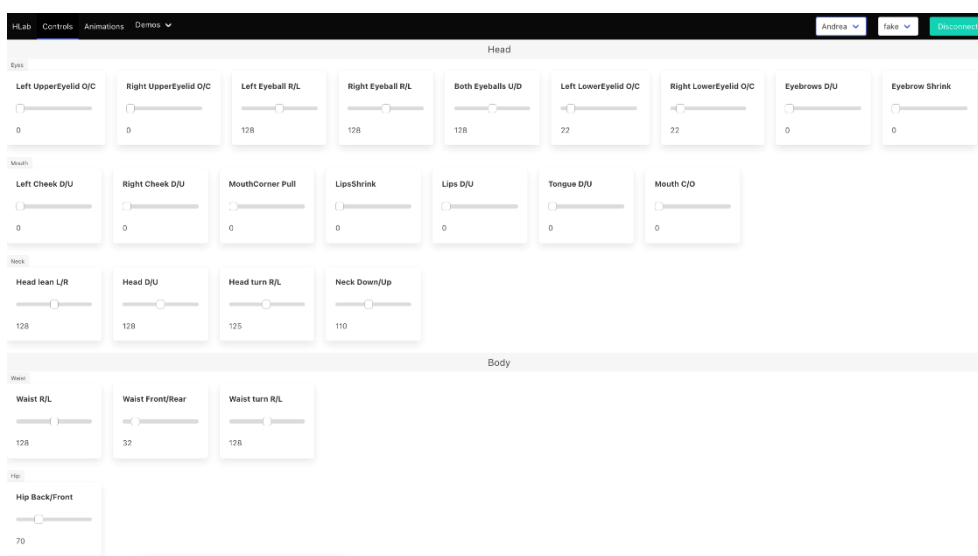
Aktuatornr.	Aktuatorbeschreibung	Bewegungsextrema
1 / 2	Linkes / rechtes, oberes Augenlid	offen / geschlossen
3 / 4	Linkes / rechtes Auge	links / rechts
5	Beide Augen	oben / unten
6 / 7	Linke / rechtes, unteres Augenlid	offen / geschlossen
8	Augenbrauen	unten / oben
9	Augenbrauen zusammenziehen	entspannt / angespannt
10 / 11	Linke / rechte Wange	unten / oben
12	Beide Mundwinkel	entspannt / angespannt
13	Beide Lippen schürzen	entspannt / angespannt
14	Beide Lippen öffnen	geschlossen / offen
15	Zunge	unten / oben
16	Mund	geschlossen / offen
17	Kopf neigen	links / rechts
18	Kopf neigen	vorne / hinten
19	Kopf drehen	rechts / links
20	Nacken neigen	vorne / hinten
21	Taille neigen	rechts / links
22	Taille neigen	vorne / hinten
23	Taille drehen	rechts / links
24	Hüfte heben	hinten / vorne
25 / 39	Linke / rechte Schulter heben	unten / oben
26 / 40	Linke / rechte Schulter drehen	hinten / vorne
27 / 41	Linken / rechten Arm drehen	hinten / vorne
28 / 42	Linken / rechten Arm drehen	unten / oben
29 / 43	Linken / rechten Oberarm drehen	im Uhrzeigersinn
30 / 44	Linken / rechten Ellenbogen drehen	ausgestreckt / angewinkelt
31 / 45	Linken / rechten Unterarm drehen	im Uhrzeigersinn
32 / 46	Linkes / rechtes Handgelenk drehen	im Uhrzeigersinn

Aktuatornr.	Aktuatorbeschreibung	Bewegungsextrema
33 / 47	Linkes / rechtes Handgelenk anheben	unten / oben
34 – 38 / 48 - 52	Jeder Finger der linken / rechten Hand einzeln	ausgestreckt / angewinkelt

Auflistung der 52 Aktuatoren des Androiden Roboters Andrea mit der jeweiligen Bewegungsbeschreibung für die Minimal- und Maximalwerte des Aktuators

Zusätzlich ist je eine kleine Kamera in die beiden Augen von Andrea integriert, deren Bilddaten über jeweils ein USB-Kabel an einen externen Computer übermittelt werden.

Über eine grafische Benutzeroberfläche, die als Webanwendung erstellt wurde und von den Mitarbeitenden des HumanoidLabs zusammen mit den Studierenden der Hochschule kontinuierlich erweitert wird, können die Bewegungsmöglichkeiten von Andrea exploriert und Animationen erstellt werden, vgl. Abbildung unten. Letztere werden im JSON-Format (Bourhis u. a. 2017) abgespeichert, bei einem Refresh des Browsers dynamisch geladen und können dann über einen Knopfdruck in der Benutzerschnittstelle oder programmatisch an- und ausgeschaltet werden.



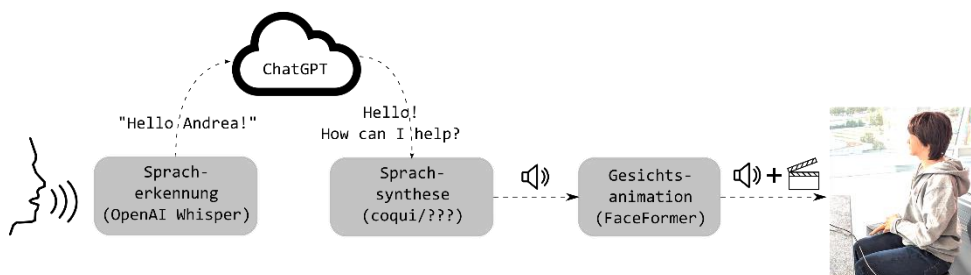
Die Benutzerschnittstelle der Webanwendung zur Kontrolle des Androiden Roboters Andrea.

Animationen enthalten auch Angaben dazu, wie oft sie automatisch wiederholt werden sollen, ob eine Pause zufälliger Länge nach jeder

Iteration eingefügt werden soll und welche Priorität die Animation hat. Damit können mehrere Animationen, die unterschiedliche Aktuatoren animieren, dynamisch parallel ausgeführt werden, wie zum Beispiel ein kontinuierliches Augenblinzeln zusammen mit einer Atemanimation des Oberkörpers.

1.2.1 Andrea als verkörperter Chatbot

Auf der bis hierher beschriebenen Basistechnologie aufbauend, wurde eine erste Integration verschiedener Methoden des maschinellen Lernens realisiert. Die Anbindung von ChatGPT 3.5 im Sommer 2023 ermöglicht es einem Interaktionspartner oder einer Interaktionspartnerin nun, sich per Mikrofon mit Andrea zu unterhalten. Die folgende Abbildung veranschaulicht den Prozess von der Spracheingabe bis zur multimodalen Sprachausgabe durch den Androiden Andrea und wird in (Heisler, Kopp, und Becker-Asano 2023) detailliert beschrieben.



Der grundsätzliche Ablauf der Antwortgenerierung und animierten Sprachausgabe durch Andrea.

Nachdem das Audiosignal der Spracheingabe vollständig aufgenommen wurde, wird es mittels des vortrainierten, künstlichen, neuronalen Netzes „Whisper“ in Text umgewandelt. Dieser Prozess wird als Spracherkennung bezeichnet. Der so erkannte Text wird mit einem vordefinierten System-Prompt ergänzt per API-Aufruf zum ChatGPT-Server gesendet. Von dort erhält das System die generierte Antwort als Text zurück. Eine Sprachsynthese, die ebenfalls auf einem vortrainierten, künstlichen, neuronalen Netz basiert, transformiert diesen Text in ein digitales Audiosignal, welches einem weiteren, künstlichen, neuronalen Netz als Eingabe zur Generierung einer passenden Gesichtsanimation dient. Schließlich wird das Audiosprachsignal synchron zu der Gesichtsanimation

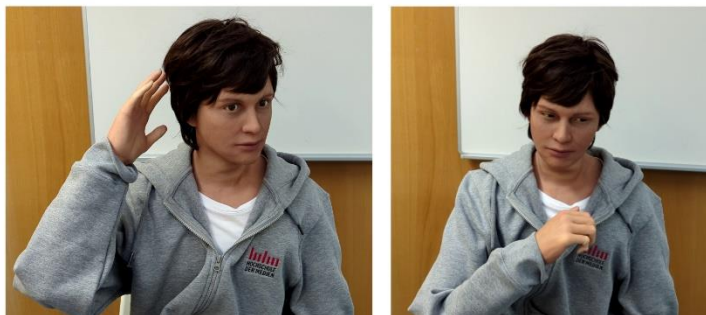
abgespielt. Es erscheint nun so, als würde Andrea natürlichsprachlich antworten.

Zur Kontrolle der Spracheingabe wurde ein Taster vor einem Mikrofon auf dem Tisch platziert. Dies ist in der folgenden Abbildung zu erkennen.



Android Andrea in der Interaktion mit drei Besuchern des Mercedes-Benz-Museums.

Sobald der Taster betätigt wird, startet eine vordefinierte Animation, durch die Andrea den Kopf zur Seite dreht und die Hand ans Ohr hebt. Gleichzeitig wird das Mikrofon aktiviert und das Audiosignal so lange aufgenommen, bis der Taster wieder losgelassen wird. Dann wird eine zweite Animation gestartet, die Andrea die Hand ans Kinn nehmen und mit halb geschlossenen Augen nach unten blicken lässt. Auch wird der Kopf immer wieder leicht hin und her geschwenkt. In der folgenden Abbildung können beide Animationen nachvollzogen werden.



Standbilder der beiden Animationen während der Audioaufnahme, links, und der Antwortgenerierung, rechts.

Nachdem alle Prozesse die Antwortgenerierung wie oben beschrieben abgeschlossen haben, hebt Andrea den Kopf und es werden die Kamerabilder des linken Auges von Andrea ausgewertet, um mittels eines weiteren, vortrainierten, neuronalen Netzes die Anwesenheit von Personen zu erkennen. Falls Personen erkannt wurden, werden Animationen dynamisch erzeugt, die den Kopf von Andrea so bewegen, dass eine dieser Personen im Videostream zentriert wird. Die Auswahl der Person erfolgt derzeit zufällig. Parallel beginnt die Audioausgabe zusammen mit der Gesichtsanimation wie weiter oben beschrieben.

1.3 Ergebnisse einer explorativen Studie zur Akzeptanz von Andrea im Museum

Das im vorigen Abschnitt beschriebene System wurde von Dienstag, 31.10.23, bis Sonntag, 5.11.23, im Bereich der Galerie der Reisen des Mercedes-Benz-Museums installiert. Eine Vorankündigung fand bewusst nicht statt, um möglichst spontane Reaktionen der Besucherinnen und Besucher beobachten zu können.

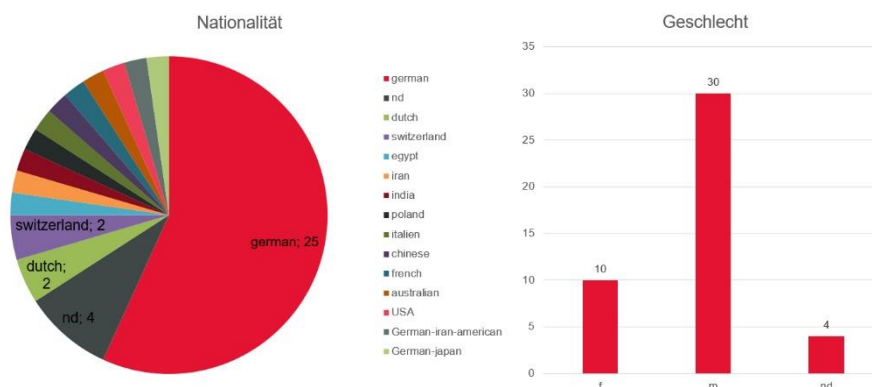
Somit ergab sich die Chance, die Wirkung von Andrea auf die Besucher und Besucherinnen qualitativ durch strukturierte Interviews zu erfassen. Diese waren wie folgt strukturiert:

1. Haben Sie mit dem Roboter interagiert?
Ja → weiter mit Fragen 1a) und 1b)
Nein → weiter mit Fragen 1c) und 1d)
 - a) Was hat Sie motiviert aktiv mit dem Roboter zu interagieren?
 - b) Beschreiben Sie bitte Ihre Interaktion mit dem Roboter.
 - c) Weshalb haben Sie nicht aktiv interagiert?
 - d) Beschreiben Sie Ihre Beobachtungen.
2. Für wie sinnvoll halten Sie den Einsatz dieses Roboters im Museum heute auf einer Skala von 0 (gar nicht sinnvoll) bis 10 (sehr sinnvoll)?
3. In welchen Situationen wäre der Einsatz im Museum NICHT sinnvoll?
Warum nicht?
4. In welchen Situationen wäre der Einsatz im Museum sinnvoll? Warum?

5. Was sollte dringend verbessert werden, bevor ein solcher Roboter eingesetzt wird?
6. Haben Sie noch Fragen und/oder Anregungen?

Eine englische Übersetzung dieses Interviewleitfadens wurde ebenfalls vorbereitet und je nachdem welche Angaben zur Herkunft vorab gemacht wurden (zusätzlich zu Altersgruppe und Geschlecht) wurde das Interview auf Englisch oder Deutsch durchgeführt.

Vierundvierzig BesucherInnen waren nach einer allgemeinen Aufklärung über das Ziel des Interviews und der Versicherung, dass die Daten lediglich anonymisiert ausgewertet würden, zum Interview bereit und unterschrieben eine Einwilligungserklärung. Die Herkunftsverteilung sowie die Geschlechterverteilung sind auf folgender Abbildung zusammengefasst.

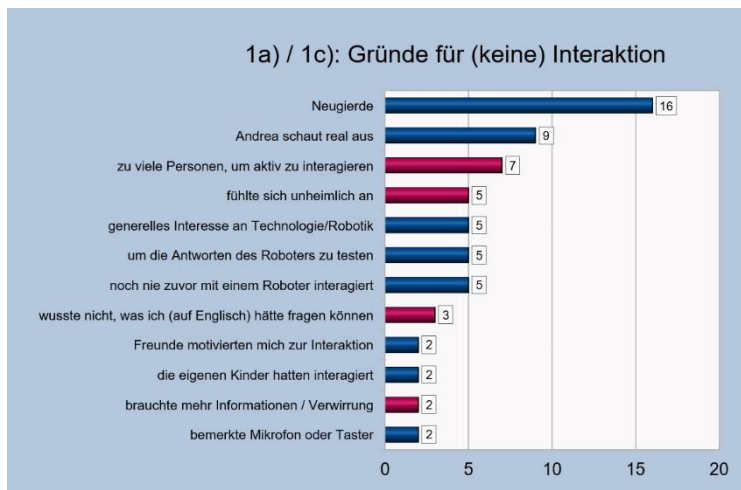


Verteilungen zu Herkunft (links) und Geschlecht (rechts) der Befragten (bei je vier Personen fehlte die Angabe zu Herkunft bzw. Geschlecht und diese sind mit „nd“ gekennzeichnet)

Die Antworten wurden mit einem Audiorekorder aufgezeichnet, dann transkribiert und die erhaltenen Texte manuell bereinigt und die Sprachanteile des/der InterviewerIn und des/der BesucherIn entsprechend annotiert. Eine erste manuelle Analyse aller Antworten wird im Folgenden berichtet.

1.3.1 Gründe für das (nicht) interagieren

Die Ergebnisse der gemeinsamen Analyse der Antworten auf die Fragen 1a) und 1c) werden im folgenden Balkendiagramm dargestellt.



Die Analyseergebnisse zu den Antworten auf die Fragen 1a) und 1c) mit ihren absoluten Häufigkeiten. Antworten zu 1a) werden durch blaue Balken dargestellt, Antworten zu 1c) durch rote. Manuell ermittelte Cluster mit nur einem Element wurden als irrelevant betrachtet und werden nicht dargestellt.

Es fällt auf, dass die Besucher und Besucherinnen vor allem aus Neugierde mit Andrea interagiert haben und weil das Äußere des Roboters als sehr real wahrgenommen wurde. Der Hauptgrund für eine Nicht-Interaktion war oft die Anwesenheit von zu viel Publikum. Andreas Anwesenheit wurde aber auch von einzelnen als unheimlich empfunden. Letzteres ist ein Indiz für eine Bestätigung der Uncanny Valley Hypothese (Becker-Asano u. a. 2010; Pütten u. a. 2014), bei der vermutet wird, dass ein zu realistisches, menschähnliches Design die Gefahr birgt abstoßend und unheimlich zu wirken. Es ist generell nicht verwunderlich, dass BesucherInnen eines Technikmuseums, wie es das Mercedes-Benz-Museum ist, auch an einem Roboter wie Andrea Interesse haben und ihn testen wollen, wie aus einigen Antworten herauszulesen ist.

1.3.2 Beschreibungen der Interaktion mit dem Roboter

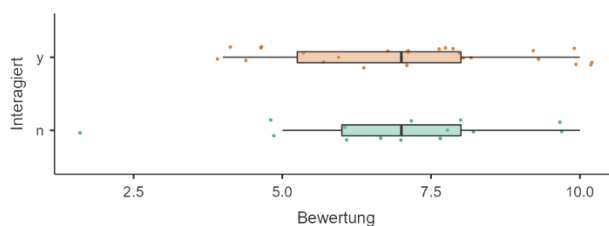
Bei der freien Beschreibung der Interaktion bzw. der passiven Beobachtungen (1b und 1d) wurde meistens lediglich berichtet, dass Fragen an Andrea gestellt wurden, wie anhand des folgenden Balkendiagramms erkennbar ist.



Ergebnisse der Auswertung der Beschreibungen, die zu Andrea geliefert wurden in Form aggregierter, absoluter Häufigkeiten.

Auch in diesem Teil des Interviews wurden vereinzelt negative Aspekte der Interaktion erwähnt, wie z.B. die nicht flüssige bzw. nicht natürliche Kommunikationsfähigkeit von Andrea. Auch ein Hinweis auf den Uncanny Valley Effekt lässt sich aus vier der Antworten herauslesen.

1.3.3 Bewertung der Sinnhaftigkeit des Einsatzes von Andrea im Museum

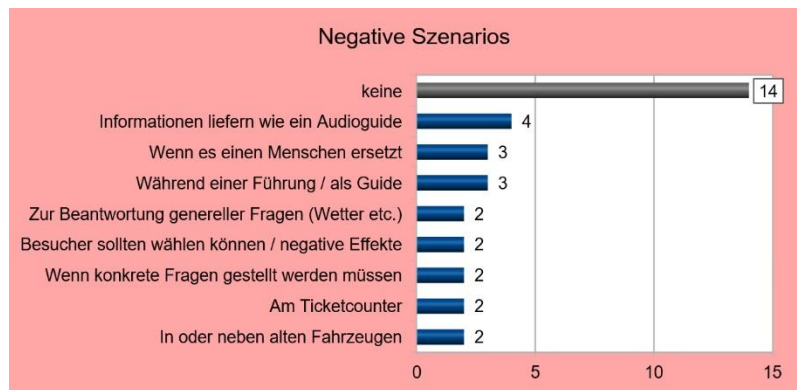


Box-plot der Daten zur Bewertung (auf einer Skala von 0 bis 10) der Sinnhaftigkeit des Einsatzes von Andrea im Museum mit Median und Perzentilen aufgeteilt nach der Beantwortung der Frage, ob vorher aktiv interagiert worden ist (oben) oder nicht (unten).

Wie in der obigen Abbildung erkennbar, wurde die Sinnhaftigkeit von beiden Personengruppen (26 interagierende und 14 nicht interagierende) relativ gleich bewertet. Der Median beider Verteilungen ist sieben und die Perzentile sind nur im unteren Bereich etwas unterschiedlich. Eine einfaktorielle ANOVA mit der Annahme ungleicher Varianzen (Welch) lässt mit einem einem p-Wert von 0,829 ($F = 0,0478$; $df_1 = 1$; $df_2 = 25,2$) ebenfalls nicht vermuten, dass die beiden Verteilungen signifikant unterschiedlich sind. Somit kann davon ausgegangen werden, dass der Einsatz von Andrea von beiden Gruppen gleich positiv bewertet wurde, denn der Median sieben liegt über dem theoretischen Durchschnittswert von fünf.

1.3.4 Wo Andrea nicht eingesetzt werden sollte

Die dritte Frage schien viele der Befragten überrascht zu haben. Trotzdem wurden Angaben zu Szenarien gemacht, in denen ein Roboter wie Andrea im Museum oder auch allgemein nicht eingesetzt werden sollte. Deren Häufigkeit ist in der folgenden Abbildung als Balkendiagramm dargestellt.

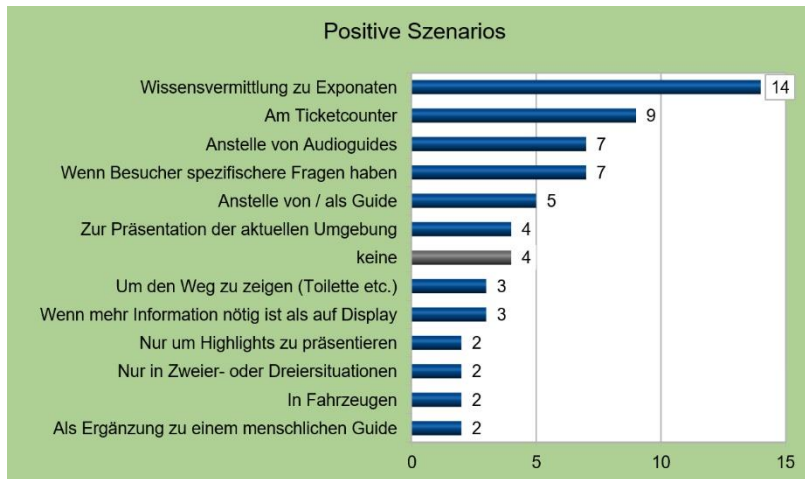


Zusammenfassung der Antworten auf die Frage, in welchen Szenarien Andrea nicht eingesetzt werden solle. All die Szenarien, die nur von einer Person genannt wurden, werden auch hier nicht aufgeführt.

Etwa jede dritte befragte Person konnte kein Szenario nennen, in dem Andrea im Museum nicht eingesetzt werden könnte. Vier Personen fänden es unnötig, wenn Andrea nur die bereits vorhandenen Audioguides ersetzen würde. Von einigen wurde es kritisch gesehen, wenn Menschen durch den Roboter ihren Job verlieren würden und eine allzu große Diskrepanz zwischen alter und neuer Technologie sollte ebenfalls vermieden werden.

1.3.5 Wo Andrea eingesetzt werden könnte

Bei der positiv gestellten Anschlussfrage fiel auf, dass ebenfalls ein Drittel der Befragten sich von Andrea Wissen über die Exponate vermitteln lassen würde. Die folgende Abbildung gibt einen Überblick.

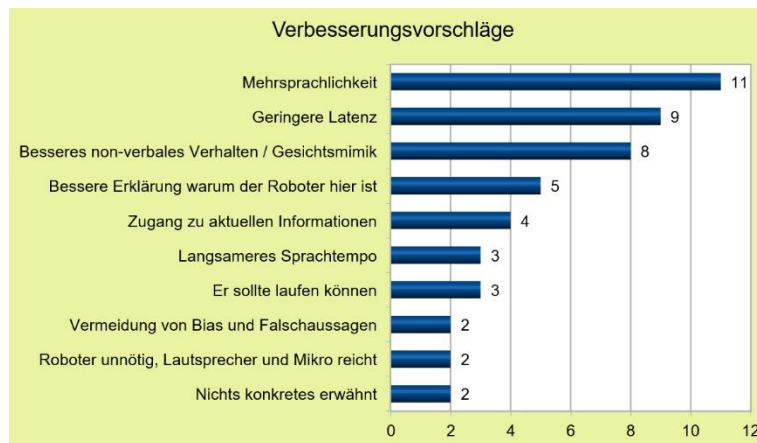


Eine Übersicht über die Häufigkeit der genannten positiven Szenarien, in denen die Besucher sich Andrea vorstellen könnten.

Zudem ist es bemerkenswert, dass einige der Befragten sich Andrea auch am Ticketcounter vorstellen konnten, obwohl dies von anderen gerade ausgeschlossen wurde. Auch als Ersatz für Audioguides wurde Andrea positiv bewertet, was erneut im Widerspruch zu Aussagen anderer Besucher steht, die dies für unnötig hielten. Lediglich vier Befragten fiel kein positives Szenario ein. Es ist zu bedenken, dass der Einsatz von Andrea in anderen, weniger technisch orientierten Umgebungen vermutlich anders bewertet würde, da dort eine andere Klientel befragt werden könnte.

1.3.6 Was an Andrea verbessert werden sollte

Bisher erwartet unser System englische Sprache als Eingabemodalität und Andrea antwortet nur auf Englisch. Wenn Deutsch gesprochen wird, dann wird dies zwar erkannt, aber von der Spracherkennungskomponente sogleich auf Englisch übersetzt, bevor es als Prompt zu ChatGPT übermittelt wird. Diese Einschränkung sollte dringend verbessert werden, wie aus der folgenden Abbildung hervorgeht.



Vorschläge zur Verbesserung des Gesamtsystems.

Obwohl Andreas Körperanimationen positiv bewertet wurden, kam oft der Hinweis, dass die Antwortzeit verbesserungswürdig sei. Auch sollte Andrea vielfältigere Animationen sowohl für die Gestik als auch für die Mimik zeigen. Eine klarere Einbettung in einen spezifischen Kontext des Museums wurde ebenso für sinnvoll gehalten wie Zugang zu aktuellen Informationen. Drei Befragte fänden es gut, wenn Andrea laufen könnte, was technisch aufgrund der extern zugeführten Luftdruckversorgung bis auf weiteres unmöglich erscheint.

1.4 Zusammenfassung und Ausblick

Einen Roboter wie Andrea in freier Wildbahn der Bevölkerung zur freien Interaktion zur Verfügung zu stellen, war in Deutschland eine Premiere. Es war keinesfalls garantiert, dass dieses Gesamtsystem täglich acht Stunden ohne menschliche Intervention dem Ansturm der Besucher und Besucherinnen standhalten würde. Nur 2,5% der Audioeingaben konnten vom System aufgrund interner Fehler nicht beantwortet werden. Umso erfreulicher ist es, dass die positiven Meinungen zu Andrea überwiegen.

Eine technische Umsetzung innerhalb nur eines Jahres wäre ohne die Anbindung eines Large Language Models wie ChatGPT 3.5 sehr wahrscheinlich unmöglich gewesen. Die geschickte Formulierung eines System-Prompts ermöglichte es uns, ChatGPT Antworten erzeugen zu lassen, anhand derer es fast immer so erschien, als würde ein Androider Roboter namens Andrea über sich selbst sprechen.

Da den Besucherinnen und Besuchern kein Hinweis gegeben wurde, warum Andrea installiert worden war, fragten sie relativ oft (36% aller Eingaben) persönliche Fragen wie z.B. „Wer bist Du?“ oder „Wie ist Dein Name?“ aber auch „Was ist Deine Lieblingsfarbe?“. In etwa 18% der Dialoge wurden Standardphrasen wie „Hallo!“, „Tschüss!“ und „Vielen Dank!“ verwendet und nur 13% der Anfragen bezogen sich konkret auf den Museumskontext.

Vereinzelt bekamen wir die Rückmeldung, dass Andrea zum Beispiel als sozialer Interaktionspartner ältere Menschen vor Vereinsamung schützen könne. Auch konnte sich ein Zahnarzt aus der Schweiz gut vorstellen, Andrea am Empfang seiner Praxis einzusetzen. Da das HumanoidLab der Hochschule der Medien zusätzlich fünf separate Roboterköpfe mit zu Andrea ähnlicher, aber einfacherer Gestalt besitzt, können diverse Aspekte der Kommunikation parallel erforscht werden. So wurde bereits die emotionale Ausdrucksfähigkeit des Gesichts der Roboter positiv evaluiert (Kassner und Becker-Asano 2023) und eine geschlechtsneutrale Stimme für Andrea ermittelt und empirisch getestet (Kuch, Melchior, und Becker-Asano 2023).

Parallel zur Entwicklung unseres Demonstrators sind inzwischen verbesserte KI-Modelle auf den Markt gekommen, die als Eingabe sowohl direkt Audio als auch zusätzlich Bilder verarbeiten können. Dieselben Modelle können ebenfalls bereits ein Audiosignal als Ausgabe bereitstellen. Damit erübrigt sich voraussichtlich die Notwendigkeit, einzelne Modelle miteinander zu verketteten. Dies wird die Verarbeitungsgeschwindigkeit erhöhen und damit die Antwortzeit verringern. Ein Nachteil dieser Realisierung wird allerdings sein, dass die gesamte Datenverarbeitung extern stattfinden wird. Unser aktuelles System läuft abgesehen von ChatGPT 3.5 komplett lokal auf einem NVIDIA Jetson Orin Entwicklerkit mit 64 GB Arbeitsspeicher. Dadurch wird vermieden, sensible Daten auf US-amerikanischen Servern verarbeiten zu lassen, was für industrielle Anwendungen von Vorteil ist.

In Zukunft werden wir weiter an der Realisierung eines besseren Kontextverständnisses für Andrea forschen. Da je nach

Anwendungsszenario andere Aspekte wichtig sind, planen wir Andrea im Rahmen von Kooperationsprojekten in diesen konkreten Situationen einzusetzen und die gesammelten Daten in Verbesserungsmaßnahmen einfließen zu lassen. Da die Anschaffungskosten für einen Androiden Roboter wie Andrea derzeit bei ca. 200.000 € liegen, stünde einer industriellen Anwendung dieser besonderen Bauform von Robotern unserer Meinung nach ebenfalls nichts entgegen.

Literatur

- Becker, Christian, Stefan Kopp, und Ipke Wachsmuth. 2007. „Why emotions should be integrated into conversational agents“. S. 49–68 in *Conversational Informatics: An Engineering Approach*, herausgegeben von T. Nishida. Wiley.
- Becker-Asano, C. 2008. „WASABI: Affect Simulation for Agents with Believable Interactivity“. PhD thesis, Faculty of Technology, University of Bielefeld.
- Becker-Asano, C., E. Meneses, N. Riesterer, J. Hué, C. Dornhege, und B. Nebel. 2014. „The hybrid agent MARCO: A Multimodal Autonomous Robotic Chess Opponent“. in *Proc. of the 2nd Intl. Conf. on Human-Agent Interaction*.
- Becker-Asano, Christian, Kohei Ogawa, Shuichi Nishio, und Hiroshi Ishiguro. 2010. „Exploring the uncanny valley with Geminoid HI-1 in a real-world application“. S. 121–28 in *IADIS Intl. Conf. Interfaces and HCI*.
- Bourhis, Pierre, Juan L. Reutter, Fernando Suárez, und Domagoj Vrgoč. 2017. „JSON: Data model, Query languages and Schema specification“. S. 123–35 in *Proceedings of the 36th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, PODS '17*. New York, NY, USA: Association for Computing Machinery.
- Brooks, Rodney A., Cynthia Breazeal, Matthew Marjanović, Brian Scassellati, und Matthew M. Williamson. 1999. „The cog project: building a humanoid robot“. S. 52–87 in *Computation for metaphors, analogy, and agents*. Berlin, Heidelberg: Springer-Verlag.
- Cassell, Justine. 2001. „Embodied conversational agents: Representation and intelligence in user interface“. *AI Magazine* 22(3):67–83.
- Dautenhahn, Kerstin. 1999. „Embodiment and interaction in socially intelligent life-like agents“. S. 102–42 in *Computation for Metaphors,*

Analogy and Agent, LNAI 1562, herausgegeben von C. L. Nehaniv. Springer.

- Fikes, Richard E., und Nils J. Nilsson. 1971. „Strips: A new approach to the application of theorem proving to problem solving“. *Artificial Intelligence* 2(3):189–208. doi: 10.1016/0004-3702(71)90010-5.
- Glas, Dylan F., Takashi Minato, Carlos T. Ishi, Tatsuya Kawahara, und Hiroshi Ishiguro. 2016. „ERICA: The ERATO Intelligent Conversational Android“. S. 22–29 in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*.
- Goleman, Daniel. 1995. *Emotional Intelligence*. New York: Bantam Books.
- Hashimoto, Kenji, und Atsuo Takanishi. 2019. „WABIAN and Other Waseda Robots“. S. 265–89 in *Humanoid Robotics: A Reference*, herausgegeben von A. Goswami und P. Vadakkepat. Dordrecht: Springer Netherlands.
- Hashimoto, S., S. Narita, H. Kasahara, und K. Shirai. 2002. „Humanoid Robots in Waseda University - Hadalz-2 and WABIAN“. S. 25–38 in *Autonomous Robots 12*.
- Heisler, Marcel, Stefan Kopp, und Christian Becker-Asano. 2023. „Making an Android Robot Head Talk“. in *Proc. of 32nd IEEE Int. Conf. on Robot and Human Interactive Communication (RO-MAN)*.
- Holodyski, Manfred, und Wolfgang Friedlmeier. 2005. *Development of Emotions and Emotion Regulation*. Springer.
- Ishiguro, Hiroshi. 2005. „Android Science: Toward a new cross-disciplinary framework“. S. 1–6 in *Proceedings of the CogSci 2005 Workshop "Toward Social Mechanisms of Android Science"*.
- Jung, Frederike, Heiko Müller, und Susanne CJ Boll. 2022. „It’s Not Warm But That’s Okay: About Robots That Avoid Human Stereotypes“. S. 1–15 in *Nordic Human-Computer Interaction Conference, NordiCHI '22*. New York, NY, USA: Association for Computing Machinery.
- Kassner, Amelie, und Christian Becker-Asano. 2023. „Comparing an android head with its digital twin regarding the dynamic expression of emotions“. in *2023 11th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*.
- Kuch, Johanna, Frank Melchior, und Christian Becker-Asano. 2023. „Effects of gender neutralization on the anthropomorphism of voices“. in *Proc. of 32nd IEEE Int. Conf. on Robot and Human Interactive Communication (RO-MAN)*.

- Lindblom, Jessica, und Tom Ziemke. 2003. „Social Situatedness of Natural and Artificial Intelligence: Vygotsky and Beyond“. *Adaptive Behavior* 11:79–96.
- Loveys, Kate, Gabrielle Sebaratnam, Mark Sagar, und Elizabeth Broadbent. 2020. „The Effect of Design Features on Relationship Quality with Embodied Conversational Agents: A Systematic Review“. *International Journal of Social Robotics* 12(6):1293–1312. doi: 10.1007/s12369-020-00680-7.
- Mainzer, Klaus. 2019a. „Von der natürlichen über die künstliche zur Superintelligenz?“ S. 185–243 in *Künstliche Intelligenz – Wann übernehmen die Maschinen?, Technik im Fokus*, herausgegeben von K. Mainzer. Berlin, Heidelberg: Springer.
- Mainzer, Klaus. 2019b. „Wie sicher ist Künstliche Intelligenz?“ S. 245–65 in *Künstliche Intelligenz – Wann übernehmen die Maschinen?, Technik im Fokus*, herausgegeben von K. Mainzer. Berlin, Heidelberg: Springer.
- Nilsson, Nils J. 2009. *The Quest for Artificial Intelligence*. Cambridge: Cambridge University Press.
- Nishio, Shuichi, Hiroshi Ishiguro, und Norihiro Hagita. 2007. „Geminoid: Teleoperated Android of an Existing Person“. S. 343–52 in *Humanoid Robots, New Developments*, herausgegeben von A. C. de P. Filho. Vienna, Austria: I-Tech.
- Ortony, Andrew, Gerald L. Clore, und Allan Collins. 1988. *The Cognitive Structure of Emotions*. Cambridge: Cambridge University Press.
- Pütten, Astrid M. Rosenthal-von der, Nicole C. Krämer, Christian Becker-Asano, Kohei Ogawa, Shuichi Nishio, und Hiroshi Ishiguro. 2014. „The Uncanny in the Wild. Analysis of Unscripted Human-Android Interaction in the Field“. *I. J. Social Robotics* 6(1):67–83. doi: 10.1007/s12369-013-0198-7.
- Russell, Stuart, und Peter Norvig. 2021. „Artificial Intelligence, Global Edition“. Abgerufen 6. Januar 2024 (<https://elibrary.pearson.de/book/99.150005/9781292401171>).
- Sietz, Hillary. 2004. „I, Robot (Movie)“.
- Speck, David, Christian Dornhege, und Wolfram Burgard. 2017. „Shakey 2016 - How Much Does it Take to Redo Shakey the Robot?“ *IEEE Robotics Autom. Lett.* 2(2):1203–9. doi: 10.1109/LRA.2017.2665694.
- Spielberg, Steven. 2001. „A.I. (Movie)“.

Standage, Tom. 2002. *Der Türke: Die Geschichte des ersten Schachautomaten und seiner abenteuerlichen Reise um die Welt.* Campus Verlag.