

BACHELORTHESIS
Leon Schwarz

Soziale Agenten für den Bereich einer modernen Wohnumgebung

FAKULTÄT TECHNIK UND INFORMATIK
Department Fahrzeugtechnik und Flugzeugbau

Faculty of Computer Science and Engineering
Department of Automotive and
Aeronautical Engineering

Betreuung durch: Prof. Dr. Kai von Luck / Susanne Draheim
Eingereicht am: 10. Mai 2021

Leon Schwarz

Title der Arbeit

Soziale Agenten für den Bereich einer modernen Wohnumgebung

Stichworte

Soziale Agenten, Companions, modernes Wohnen, Living Place

Kurzzusammenfassung

Diese Arbeit stellt die Grundlagen des Themenbereichs sozialer Agenten im Anwendungsgebiet moderner Wohnumgebungen vor. Die Absicht ist es somit, einen Einstieg in das forschende Lernen an der HAW-Hamburg zu schaffen. Mittels der Einordnung von Begrifflichkeiten, der Vorstellung von gängigen Dienstfeldern und einer Marktübersicht wird sich der Frage genähert, was das Themengebiet der sozialen Agenten umfasst. Anschließend wird mit allgemeinen Designüberlegungen zu der Frage nach den Umsetzungsweisen soziale Agenten aus technischer Sicht übergeleitet. Es werden die Kernfunktionen eines sozialen Agenten vorgestellt und Ansätze für die Architektur sozialer Agenten beschrieben. Abschließend folgen Designempfehlungen, welche aus den vorgestellten Aspekten geschlussfolgert werden.

Leon Schwarz

Title of the Thesis

Social Agents for the Area of a modern Living Environment

Keywords

Social Agents, Companions, modern Living, Living Place

Abstract

This work presents the basics of the field of Social Agents in the application area of modern living environments. The intention is hereby, to create an entry for the exploitative learning at the HAW-Hamburg. Through the categorization of terminologies, the presentation of popular service areas and a market overview shall the question be approached, what the topic area of Social Agents includes. Subsequently general thought on the design lead over to the question of ways for the implementation of Social Agents from the technical view. The core functions of a Social Agent are presented and approaches for the architecture of Social Agents are described. Based on all prior considerations, suggestions for the design are made.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Geleitwort	1
1.2	Motivation	2
1.3	Ziel dieser Arbeit	3
2	Themenübersicht sozialer Agenten	6
2.1	Begriffsklärung	6
2.2	Dienstfelder sozialer Agenten	9
2.3	Marktübersicht	10
3	Designüberlegungen	16
3.1	Allgemeines Designvorgehen	16
3.2	Kommunikationsgrundlagen	16
3.3	Beziehungsaufbau	18
3.4	Charakter	20
3.5	Äußere Erscheinung	21
4	Kernfunktionen sozialer Agenten	24
4.1	Akustische Wahrnehmung	24
4.1.1	Sprachverständnis	25
4.1.2	Emotionserkennung	26
4.1.3	Sprecheridentifikation	27
4.1.4	Sprecherlokalisation	27
4.2	Optische Wahrnehmung	27
4.2.1	Positionserkennung	28
4.2.2	Personenidentifikation	28
4.2.3	Mimik	29
4.2.4	Gestik	29
4.2.5	Körperstellung	30
4.3	Akustische Handlung	30
4.3.1	Statische Lautsynthese	31
4.3.2	Dynamische Lautsynthese	31
4.3.3	Emotionsbehaftete Lautsynthese	32
4.3.4	Der Kommunikationsbeginn	32

4.4	Optische Handlung	33
4.4.1	Mimik	33
4.4.2	Gestik	34
4.4.3	Körperstellung	34
4.5	Berührungen	35
4.5.1	Herausforderungen	35
4.5.2	Aktive Berührungen	36
4.5.3	Passive Berührungen	36
5	Architekturen sozialer Agenten	38
5.1	Allgemeiner Aufbau	38
5.2	Prozessmodelle	39
5.3	Beispielarchitektur	42
6	Designempfehlungen	46
7	Fazit und Ausblick	49
	Abbildungsverzeichnis	52
	Referenzen	52
	Selbstständigkeitserklärung	63

1 Einleitung

1.1 Geleitwort

Der Mensch ist ein soziales Wesen. Sein Leben wird geprägt von der Gesellschaft, die ihn umgibt. Maschinen haben dabei nur eine untergeordnete Rolle, da es ihnen an sozialen Fähigkeiten mangelt. Bis heute kann der Mensch daher, bis auf wenige Ausnahmen, unterscheiden, ob es sich bei seinem Interaktionspartner um ein Lebewesen oder eine Maschine handelt. Erscheinungsbild, Verhaltens- und Umgangsformen zeichnen ihm eine klare Linie zwischen Lebewesen und Maschine auf. Diese Linie auszuradieren ist der Kernaspekt des Themenbereichs sozialer Agenten. Der perfekte soziale Agent ist eine Maschine, die sich nicht mehr von einem Menschen unterscheiden lässt.

Für den Einsatz von Maschinen im Bereich der sozialen Agenten ergeben sich dabei völlig neue Möglichkeiten. Nicht nur kann die Bedienung dieser Maschinen wesentlich einfacher gestaltet werden, es erschließen sich für diese auch völlig neue Einsatzgebiete. Soziale Fähigkeiten erlauben es Maschinen sich nahtlos in die Gesellschaft und in das soziale Leben zu integrieren. Weiterhin erhalten die Maschinen über ihre sozialen Fähigkeiten das Werkzeug, um nicht nur mit toter Materie, sondern auch mit Menschen und Tieren arbeiten zu können.

Unser Verständnis von Maschinen muss sich daher deutlich wandeln. Die Maschine wird vom stumpfen Befehlsempfänger zum Freund, zum Helfer oder zum Arbeitskollegen. Auch unsere Definition von Arbeit wird sich durch das Ausstatten der Maschinen mit sozialen Fähigkeiten verändern. Der Fokus für menschliche Arbeit wird sich immer mehr auf das Generieren von abstrakten Ideen richten, während ausführende Arbeiten immer weiter durch Maschinen mit sozialen Fähigkeiten abgelöst werden. Aus „Do what I say“ wird „Do what I mean“.

Die Auswirkung der sozialen Agenten auf den Menschen ist dabei weitestgehend unklar. Neben vielen Erleichterungen und Vorteilen bringt diese Entwicklung allerhand Herausforderungen mit sich. Auf der einen Seite steht eine neue Ebene der Dienstleistung und der Arbeit, auf der anderen Seite stehen Fragen nach Ethik, gesellschaftlichen Auswirkungen und Gesetzgebung.

Das Feld der sozialen Agenten existiert schon sehr lange. Durch den technischen Fortschritt in den letzten fünfzehn Jahren hat sich deren Entwicklung jedoch stark beschleunigt.

nigt. Immer mehr Meilensteine auf dem Weg zur sozialen Maschine werden erreicht, was das Interesse an diesem Themenbereich fördert.

1.2 Motivation

Diese Arbeit ist für die Forschungseinrichtung Living Place der HAW-Hamburg geschrieben. Das Living Place dient primär der Erforschung des zukünftigen Wohnens. Eine ausführlichere Vorstellung des Living Places kann in Abschnitt 1.3 gefunden werden. Aktuell befindet sich ein Forschungszweig in Richtung sozialer Agenten im Aufbau. Die Motivation für die Forschung an sozialen Agenten ist dabei vielfältig.

Allgemeine Motive

Allgemein ist zu sagen, dass die Forschung an sozialen Agenten gleich unter mehreren Aspekten lohnenswert ist. Zum einen sind die Auswirkungen von sozialen Agenten auf die Gesellschaft und das Selbstverständnis des Menschen enorm, zum anderen ist das Themengebiet der sozialen Agenten ein sehr multidisziplinäres Feld und besitzt damit eine große Zielgruppe.

Die Auswirkungen sozialer Agenten werden zurzeit sehr kontrovers diskutiert. Zu den positiven Aspekten zählen der wahrscheinliche Evolutionsschritt in der Arbeitswelt, die einfachere Gestaltbarkeit des sozialen Umfelds und die Steigerung der Autonomie des Individuums. Zu den negativen Aspekten werden die für die Interaktion mit dem Menschen notwendigen Fähigkeiten zur Manipulation und Täuschung des Menschen, der weitere Verlust an menschlicher Kompetenz und das weitere Eindringen in die Privatsphäre des Individuums genannt. Weiterhin ist zu bedenken, dass der Mensch, sobald der Agent nicht mehr von diesem unterscheidbar ist, seine Einzigartigkeit verliert. Mit entsprechenden Ressourcen kann jeder nach Belieben Menschen designen und produzieren. Im Bereich der sozialen Netzwerke ist diese Zukunft mittels sogenannter SocialBots schon Realität. Die Geschichte des Frankenstein wird durch soziale Agenten zur Wirklichkeit.

Für die Entwicklung sozialer Agenten benötigt es dabei mindestens Fachwissen aus drei Disziplinen. Zuerst muss Wissen über den Interaktionspartner, sprich den Menschen bekannt sein. Dies ist der Bereich der Psychologie, wobei besonders die Allgemein- und Sozialpsychologie relevant sind. Auf Grundlage dessen muss ein Konzept für ein neues

Wesen geschaffen werden. Hier werden Kompetenzen aus dem Charakterdesign und gegebenenfalls auch Animationsdesign benötigt. Diese Konzepte umsetzen zu können, ist Aufgabe der Technik. Schwerpunktmäßig die Informatik, aber auch Elektrotechnik und gegebenenfalls der Maschinenbau sind hierfür erforderlich.

Weiterhin ist zu überlegen, ob die Resultate der Maschinenethik in die Entwicklung sozialer Agenten miteinbezogen werden können. Dieser Teilbereich der Ethik beschäftigt sich explizit mit den Fähigkeiten, Eigenschaften und Freiheiten, die autonome Maschinen besitzen sollten.

Das Interesse weiterer Fachdisziplinen erschließt sich mit der Forschung an sozialen Agenten. Aufgrund der großen Auswirkungen auf den Menschen und die Gesellschaft, aber auch durch die neuen Perspektiven auf das Menschsein und die Arbeitswelt können viele Forschungsbereiche die Forschung an sozialen Agenten nicht ignorieren. Prominente Beispiele sind hierfür die Bereiche der Philosophie, der Soziologie und der Juristerei.

Spezifische Motive für diese Arbeit

In Bezug auf die HAW-Hamburg spricht noch ein weiterer Faktor für das Entwickeln und Forschen an sozialen Agenten. Die Arbeit an sozialen Agenten bietet der Hochschule die Möglichkeit, ein weites Spektrum an Fakultäten, Professoren und vor allem Studenten zu bedienen. Das multidisziplinäre Feld der sozialen Agenten bildet einen gemeinsamen Nenner für studiengangs- und auch fakultätsübergreifende Zusammenarbeit.

In Bezug auf die Forschung im Living Place ist das Thema der sozialen Agenten sogar zentral. Die Erforschung des zukünftigen Wohnens ist ohne das Miteinbeziehen sozialer Agenten nur eingeschränkt möglich, da die Integration sozialer Maschinen in den Privatwohnbereich schon heutzutage erfolgt und dessen Ausbau abzusehen ist.

1.3 Ziel dieser Arbeit

Diese Arbeit verfolgt das Ziel, eine Einführung in das Themengebiet sozialer Agenten zu geben und eine Orientierung für die technische Umsetzung dieser zu bieten. Die Absicht ist, auf diese Weise einen Einstieg für das forschende Lernen an sozialen Agenten im Living Places zu schaffen. Die konkrete Forschungsfrage lautet: Was beinhaltet das Themenfeld der sozialen Agenten und welche Aspekte sind bei der Umsetzung dieser in näherer Zukunft relevant?

Um diese zweiteilige Forschungsfrage beantworten zu können, wird auch die Arbeit inhaltlich in zwei Bereiche unterteilt. Der erste Teil der Forschungsfrage wird behandelt, indem verschiedene Begriffe des Themengebietes sozialer Agenten eingeordnet und häufige Aufgabenfelder aufgezeigt werden. Anschließend folgt eine Vorstellung ausgewählter Modelle auf dem Markt und Überlegungen zum allgemeinen Design der Agenten-Mensch Interaktion. Mit diesen Überlegungen leitet die Arbeit zur Beantwortung des zweiten Teils der Forschungsfrage über. Es werden nun die Kernfunktionen sozialer Agenten aus technischer Sicht beleuchtet und Aspekte der Gesamtarchitektur sozialer Agenten besprochen. Abschließend erfolgen Empfehlungen für die Umsetzung sozialer Agenten im Rahmen des Living Place auf Grundlage der vorangegangenen Betrachtungen. Die einzelnen Themen werden dabei nur in der für die Diskussion notwendigen Tiefe vorgestellt. Für ausführlichere Informationen sei auf die Referenzen der einzelnen Kapitel verwiesen.

Das Themengebiet der sozialen Agenten ist ein sehr weites Feld. Um diese Arbeit in einem angemessenen Rahmen zu halten, sind Abgrenzungen notwendig. Folglich wird sich in dieser Arbeit auf soziale Agenten im Bereich von modernen, einräumigen und vollüberdachten Privathaushalten beschränkt. Weiterhin wird sich bei der Betrachtung der Kernfähigkeiten sozialer Agenten von der Interaktion mit dem Menschen auf die Kommunikation mit dem Menschen eingeschränkt. Auch ist zu beachten, dass diese Arbeit konzeptunspezifisch erstellt ist. Es werden keine Konzepte für soziale Agenten entwickelt und die thematisierten Aspekte allgemein gehalten.

Das Living Place

Das Living Place [49] ist ein Labor der Forschungseinrichtung Creative Space for Technical Innovations (CSTI) [48] der HAW-Hamburg. Es bietet eine 140m² große, vollingerichtete Einzimmerwohnung für die Erforschung des Lebens von morgen.

Der Wohnbereich des Living Place teilt sich in Schlafzimmer, Esszimmer, Küche, Wohnzimmer und Bad auf. Über das Hausnetzwerk sind verschiedenste Smart Home Devices in die Wohnung integriert. Weiterhin verfügt das Living Place über eine Fülle von Kameras und Mikrofonen und ist mit einem Indoor Positioning System (IPS) ausgestattet.

Neben der Möglichkeit zur Datenaufnahme besitzt das Living Place verschiedenste Steuerungsmöglichkeiten. Zu den aktuellen Funktionen zählen hierbei die Steuerung von Farbe

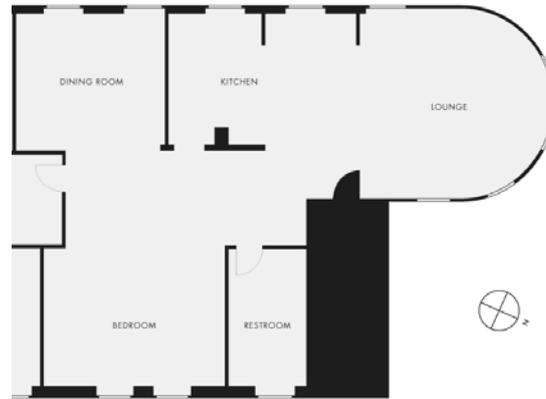


Abbildung 1: Grundriss des Laborbereiches des Living Places

und Helligkeit sämtlicher Raumbelichtungen, das Regeln der Heizung und das Öffnen und Schließen der Fenster sowie der Rollos.

2 Themenübersicht sozialer Agenten

Um zu erfassen, was das Themengebiet der sozialen Agenten beinhaltet, werden in diesem Kapitel drei Betrachtungen angestellt. Zuerst werden häufig auftretende Bezeichnungen von sozialen Agenten eingeordnet. Anschließend werden populäre Anwendungsgebiete sozialer Agenten vorgestellt. Zuletzt wird eine Marktübersicht über verschiedene soziale Agenten mit Eignung für den Einsatz in modernen Wohnräumen geboten.

2.1 Begriffsklärung

Dieses Unterkapitel definiert zuerst den Begriff sozialer Agent und ordnet daraufhin weitere Begrifflichkeiten zu den Unterkategorien sozialer Agenten ein.

Definition sozialer Agenten

Auf die Frage, was einen sozialen Agenten, beziehungsweise einen sozialen technischen Agenten, ausmacht, gibt es bis heute keine feststehende Antwort. Das Wort Agent kommt aus dem Lateinischen und steht für Treibender, Handelnder oder auch für Vermittelnder. Es werden verschiedene Begriffe und Definitionen genutzt, um den Bereich der sozialen Agenten abstecken zu können. Der Begriff Companion oder Companion System wird heutzutage oft synonym genutzt [15]. Um dieser Arbeit einen Rahmen zu geben, wird sich an zwei populären Definitionen orientiert. Zum einen ist dies die Definition des Begriffes „Agent“ nach dem VDI und zum anderen die Definition des Begriffes „*sociable Machine*“ nach Cynthia Breazeal. Der Begriff Maschine ist hierbei, wie auch in der gesamten Arbeit, als Bezeichnung für einen künstlich geschaffenen Akteur zu verstehen und nicht als aufgabenspezifischer Bewegungsapparat, wie es die ISO [29] vorsieht.

Nach dem VDI ergibt sich für einen Agenten die folgende Definition:

„Ein technischer Agent ist eine abgrenzbare (Hardware- oder/und Software-) Einheit mit definierten Zielen. Ein technischer Agent ist bestrebt, diese Ziele durch selbstständiges Verhalten zu erreichen und interagiert dabei mit seiner Umgebung und anderen Agenten.“
[86]

Das grundlegende Agentenkonzept ist hierbei ein alleinstehender Agent, welcher auf direktem Wege mit seiner Umgebung interagiert. Die Kernfunktionen dieses Agenten sind

dabei auf die Wahrnehmung seiner Umgebung und das Handeln in ihr beschränkt. Der Ablauf der Funktionen eines Agenten ist dabei repetitiv nach dem Schema „wahrnehmen, verarbeiten, handeln“ aufgebaut.

Die zweite Definition wurde von der Forscherin des MIT, Cynthia Breazeal, in dem Buch „*Socially Intelligent Agents*“ [17] veröffentlicht. Diese bildet eine der ältesten gängigen Charakterisierungen sozialer Maschinen. Hierin definiert Breazeal, dass soziale Maschinen in der Lage sind, auf persönlicher Ebene mit Menschen zu interagieren, zu kommunizieren und eine soziale Beziehung mit ihnen aufzubauen. Dazu ist die Maschine fähig, sich selbst und auch den Menschen aus sozialer Sicht zu verstehen. Gleichzeitig kann sich die Maschine auf sozialer Ebene dem Menschen gegenüber verständlich machen. Breazeal hebt hier die Fähigkeit zum Aufbau von gegenseitiger Empathie hervor. Abschließend fasst sie ihre Definition zusammen mit dem Besitz der menschlichen sozialen Intelligenz.

In weiteren Definitionen wird der hiermit abgesteckte Bereich um den intuitiven Umgang oder die Lernfähigkeit der sozialen Maschine ergänzt. Häufig wird ebenfalls der Begriff des Always-On Systems mit der Definition eines sozialen Agenten verbunden. Always-On Systeme gehören hierbei einem eigenen Themenbereich an, welcher sich mit dem Verhalten von Maschinen auseinandersetzt, die vom Nutzer nicht aktiv an- oder ausgeschaltet werden [74].

Einordnung weiterer Begrifflichkeiten

Innerhalb des Definitionsrahmens sozialer Agenten, werden eine Fülle verschiedenster Begriffe für die Bezeichnung von Untergruppen verwendet. Viele Konzepte gehen dabei über die Fähigkeiten eines sozialen Agenten hinaus, sodass ihre sozialen Fähigkeiten nur Teilfunktionen darstellen. Um einen Überblick über den aktuellen Stand zu geben, werden die bekanntesten Begrifflichkeiten hier vorgestellt.

Zum einem werden soziale Agenten nach ihrem Aufbau bezeichnet. Ein großer Bereich ist hier das Feld der Social Robots. Einfach gesagt, sind dies soziale Agenten, welche über einen bewegungsfähigen Körper verfügen [36]. Auch viele humanoide Roboter können mit zu den sozialen Roboter gezählt werden. Humanoid bedeutet dabei menschenähnlich, wobei sich dies laut Definition nur auf das Erscheinungsbild und die Bewegungsfähigkeit des Roboters beschränkt [23]. Da Menschenähnlichkeit jedoch ein weitumfassender Begriff ist, bleibt offen, ob und wie diese Bezeichnung auch andere Merkmale einer, auch körperlosen, Maschine charakterisieren könnte.

Zum anderen gibt es verschiedenste Bezeichnungen von Nachbauten verschiedener Aspekte eines Lebewesens, hinter welchen sich ein sozialer Agent verbergen kann. Artificial Pet oder Virtual Friend sind Beispiele hierfür.

Weiterhin gibt es Bezeichnungen nach den Kernfunktionen des sozialen Agenten. SocialBots, ChatBots, VoiceBots, CoBots und CareBots sind hier gängige Begriffe. Diese seien im Folgenden vorgestellt.

- **SocialBots**

SocialBots dürfen nicht mit Social Robots verwechselt werden. Der Begriff SocialBot meint ausschließlich Software, die Useraccounts auf sozialen Netzwerken betreibt [37]. Ihre Aufgabenfeld beschränkt sich auf das Werben, Ausspionieren und Manipulieren der menschlichen Nutzer innerhalb dieser Netzwerke.

- **ChatBots und VoiceBots**

Weniger spezifische Aufgabenfelder fallen ChatBots und VoiceBots zu. Sie sind einzig durch ihre Fähigkeit definiert, in der menschlichen Sprache kommunizieren zu können. Der ChatBot vermag dies in der Textform und der VoiceBot in gesprochenen Lauten [7].

- **CoBots**

Der Begriff CoBots ist eine Abkürzung der Bezeichnung Kollaborationsroboter. Wie bereits der Name verrät, ist der CoBot durch seine Fähigkeit mit dem Menschen zusammenzuarbeiten definiert [20].

- **CareBots**

Noch einen Schritt weiter in der Interaktion mit dem Menschen geht der CareBot. Nicht die Zusammenarbeit, sondern die Arbeit am Menschen selbst, also dessen Unterstützung oder Pflege, steht bei diesen sozialen Agenten als charakterisierende Fähigkeit im Vordergrund [40].

Es ist durchaus denkbar, dass eines Tages nahezu jede Maschine soziale Kompetenzen besitzen wird. Ein Zuwachs in der Anzahl der Begrifflichkeiten und das weitere Verschwimmen der Definitionen ist daher zu erwarten.

2.2 Dienstfelder sozialer Agenten

Dieses Unterkapitel stellt populäre Dienstfelder sozialer Agenten vor. Allgemein ist zu sagen, dass Soziale Agenten eine Vielzahl unterschiedlicher Dienste ausführen können. Dabei ist eine Trennung der sozialen Agenten nach Dienstfeldern nur selten möglich. Ein Dienst kann meist mehreren Dienstfeldern zugeordnet werden. Um einen besseren Überblick zu verschaffen, werden nachfolgend die gängigsten Dienstfelder sozialer Agenten vorgestellt.

- **Companionship**

Companionship, oder auch Gesellschaft, ist die reinsten Anwendungsform des sozialen Agenten. Hier steht die soziale Beziehung zu den Menschen im Vordergrund. Ohne weitere Dienstfelder ist ein sinnvolles Einsatzgebiet für diese sozialen Agenten das Ausgleichen fehlender zwischenmenschlicher Beziehungen und somit die Bekämpfung von Vereinsamung.

- **Education**

Ein populäres Dienstfeld sozialer Agenten ist das Lehren oder Coachen. Diese Agenten werden auch Pedagogical Agents [42] genannt. Verglichen zum Menschen bieten sie den Vorteil, dass sie sich in verschiedenster Hinsicht besser auf die Lernenden einstellen können. Bei der Arbeit mit Kindern ist es ihnen beispielsweise möglich als Gleichgesinnter aufzutreten, womit dessen Akzeptanz im Vergleich zu einem lehrenden Erwachsenen wesentlich höher sein kann [19]. Die Akzeptanz des sozialen Agenten bietet damit die Möglichkeit einer effizienteren Lehrform. Anstelle des Frontalunterrichts kann der Agent die Illusion schaffen, als Freund zusammen mit dem Lernenden das Themenfeld zu entdecken. Dies gilt auch im Bereich des Sports, wo der soziale Agent als Trainingspartner auftreten kann. Weiterhin eignet sich der soziale Agent speziell für das Lehren oder Coachen im Themenbereich der sozialen Interaktion. Sei es beim Training für ein Bewerbungsgespräch oder beim Lernen einer neuen Sprache, der Agent kann jeden sozialen Interaktionspartner simulieren und somit den Umgang mit diesen üben [66].

- **Entertainment**

Ein potenziell sehr großes Dienstfeld ist der Einsatz sozialer Agenten in der Unterhaltungselektronik. Ein Anwendungsbeispiel ist die Ausstattung von Computerspielcharakteren mit sozialer Intelligenz. Was den Bereich der sozialen Roboter angeht, sind die Anschaffungskosten für Privathaushalte im Verhältnis zu deren

Fähigkeiten meist noch sehr hoch [1], weswegen diese zurzeit noch ein Luxusprodukt darstellen. Um diese Kosten auszugleichen, wird das Funktionsspektrum des sozialen Roboters meist mit Anwendungen aus weiteren Dienstfeldern vergrößert.

- **Assistent**

Das am weitesten in Privathaushalten verbreitete Dienstfeld sozialer Agenten ist die Assistenz. Alexa, Siri und Cortana sind hier die Vorreiter. Aktuell ziehen diese sozialen Agenten ihre Nützlichkeit dabei weniger aus der sozialen Beziehung, als mehr aus dem Verstehen und Ausführen direkter Anweisungen, sowie dem Speichern und Verarbeiten personenbezogener Daten. Der Raum für tiefere soziale Beziehungen, etwa im Sinne eines Butlers, ist hier jedoch vorhanden [41]. Auch als Front-End sozialer Interaktionen kann der Agent Verwendung finden. Im Namen des Nutzers kann er Texte schreiben und Telefonate führen und ersetzt somit eine einfache Servicekraft oder einen Sekretär. Der VoiceBot Google Duplex ist ein Beispiel für diese Funktionen.

- **Care**

Die Pflege ist ein weiteres Dienstfeld für soziale Agenten, an dem intensiv geforscht wird. Man verspricht sich von den Fortschritten auf diesem Feld zum einen die Entlastung von Pflegepersonal und zum anderen den längeren Erhalt der Autonomie von Menschen mit Einschränkungen [40]. Trotz einer Vielzahl von Prototypen und Forschungsgruppen gibt es aktuell noch wenige soziale Agenten, die dieses Ziel in der Praxis erreichen. Einen bescheidenen Erfolg haben soziale Agenten im Bereich der Pflege nur in Nischenanwendungen oder als Lehrer oder Coach im Umgang mit gesundheitlichen Herausforderungen.

- **Forschung**

Die meisten Modelle sozialer Agenten sind heutzutage reine Forschungsobjekte. Wie angesprochen liegt das Forschen an und mit sozialen Agenten im Interessensbereich vieler Disziplinen und ist damit entsprechend relevant. Einen großen Bereich bildet hierbei die Nachbildung von Teilaspekten des Menschen durch die Entwicklung menschenähnlicher Agenten [35].

2.3 Marktübersicht

Dieses Unterkapitel stellt eine Auswahl, für das Living Place interessanter, sozialer Agenten auf dem Markt vor. Das Spektrum an bereits existierenden sozialen Agenten ist gewaltig.

Auch wenn sich bei deren Anwendungsbereich auf einen Einpersonenhaushalt beschränkt wird, übertritt die Masse an Modellen den Rahmen dieser Arbeit deutlich. Daher werden im Folgenden nur eine Auswahl an sozialen Agenten in ihren Grundzügen vorgestellt. Dabei wird zuerst im Living Place einsetzbare Plattformen präsentiert und anschließend auf für das Living Place interessante Produkte eingegangen.

Plattformen

Wann eine Plattform als sozialer Agent bezeichnet werden kann, hängt von dessen individueller Anwendung ab. Es werden daher Plattformen vorgestellt, die sich gut für den Einsatz als soziale Agenten eignen, jedoch nicht auf diesen beschränkt sind.

Nao



Abbildung 2: NAO

Nao ist ein knapp 60 cm großer humanoider Roboter. Neben seiner handlichen Größe zeichnet ihn das Vorhandensein eines vollbeweglichen und vollständigen menschlichen Körpers aus. Weiterhin verfügt Nao über die Fähigkeit Objekte zu erkennen, Sprecher zu identifizieren sowie mehrere Sprachen zu verstehen und zu sprechen [75]. Mittels dem SDK Choregraphe ist die Inbetriebnahme von Nao durch eine graphische Programmierung möglich. Gleichzeitig kann der Roboter jedoch auch in einer Vielzahl von Programmiersprachen programmiert werden. Nao eignet sich besonders für die Forschung an Körpersprache oder auch als Bewegungscoach. Diese Plattform für das Living Place zu verwenden, bietet sich an, da das CSTI bereits über mehrere Exemplare verfügt.

Pepper



Abbildung 3: Pepper

Wie Nao wird auch Pepper zu den humanoiden Robotern gezählt. Sie ist 120 cm groß und besitzt einen vollständig beweglichen menschlichen Oberkörper. Auch sie verfügt über die Fähigkeit Objekte zu erkennen und Sprachen zu sprechen [76], und kann über Choregraphie in Betrieb genommen werden. Aufgrund ihrer Größe und ihres Aufbaus eignet sich Pepper besonders für die Interaktion mit Menschen im offenen Raum.

Furhat



Abbildung 4: Furhat

Furhat ist recht neu auf dem Markt und fokussiert in seinem Design explizit die Kommunikation mit Menschen. Er verfügt über einen beweglichen Kopf und Halsbereich und erzeugt mittels eines Projektors das Bild eines Gesichtes auf einer Silhouette aus Kunststoff. Das Gesicht kann dabei beliebig gestaltet werden und die Silhouette ist austauschbar. Furhat spricht verschiedene Sprachen und besitzt die Fähigkeit Gesichter zu erkennen, zu tracken und zu analysieren [66]. Für die Programmierung stellt der Hersteller das Furhat SDK bereit, mit welcher der Roboter graphisch programmiert werden kann.

Produkte

Im Folgenden werden soziale Agenten vorgestellt, die sich entweder in die Forschung im Living Place integrieren lassen oder interessante Beispiele für die Rubrik der sozialen Agenten als Endverbraucherprodukte darstellen.

Googles Meena und Duplex

Googles Meena [64] ist nach eigenen Angaben der am weitesten entwickelte Chatbot auf dem Markt. Meenas Kommunikationsfähigkeiten fallen in die Kategorien multi-turn und open-domain. Das bedeutet, dass sie einen Chat auch über mehrere Zyklen und zu jedem beliebigen Thema souverän führen kann. Weiterhin hat Google einen VoiceBot mit dem Namen Google Duplex entwickelt. Auch Google Duplex [46] verfügt über die Fähigkeiten des multi-turns und des open-domains. Der VoiceBot kann selbstständig Telefonate führen und ist in seiner verbalen Erscheinung nicht mehr von einem Menschen unterscheidbar. Meena ist aktuell noch nicht zur freien Verwendung verfügbar, während Duplex bereits auf den neuesten Androidgenerationen in den USA läuft. Es ist zu erwarten, dass sowohl die Verfügbarkeit als auch der Funktionsumfang in den nächsten Jahren stark wachsen wird. Die Fähigkeiten lassen sich dann möglicherweise als Plattform für Forschungsprojekte nutzen, womit diese Produkte perspektivisch auch für das Living Place von Nutzen sein könnte.

Vector und Jibo

Vector [39] und Jibo [38] sind soziale Roboter, die in den letzten Jahren als Unterhaltungsprodukt auf dem Verbrauchermarkt vorzufinden waren. Sie werden hier als Beispiele



Abbildung 5: Jibo



Abbildung 6: Vector

für das Versagen sozialer Roboter auf dem Markt vorgestellt. Beide Produkte sind inzwischen eingestellt. Der Grund für ihr Scheitern ist, neben dem hohen Anschaffungspreis, die mangelnde Fähigkeit, eine Beziehung mit einem Menschen aufrecht zu erhalten. Ihr Verhalten wird als zu repetitiv und teilweise zu fehlerhaft beschrieben, um das Interesse und die Gunst ihrer Eigentümer über längeren Zeitraum zu erhalten [85].

Alexa, Siri und Cortana

Sehr erfolgreich hingegen sind die Sprachassistenzsysteme von Amazon, Apple und Microsoft. Alexa, Siri und Cortana kämpfen um die Marktführerschaft im aktuell erfolgreichsten Bereich der sozialen Agenten. Die Hersteller scheinen bei diesen Produkten sehr darauf bedacht zu sein, die bei Jibo und Vector gemachten Fehler zu meiden, indem sie den Funktionsumfang und den Beziehungsaufbau zu dem Nutzer gering halten. In Bezug auf das Living Place könnten diese Produkte als Front-End für verschiedenste Anwendungen genutzt werden.

Gatebox

Zum Schluss dieses Kapitels sein noch eine Kuriosität vorgestellt. Ein Avatar soll hier als Hologramm die partnerschaftliche Beziehung mit einem Menschen ersetzen. Durch die Integrationsfähigkeit in das Smart Home und das Handy des Benutzers soll die Illusion eines partnerschaftlichen Zusammenlebens entstehen [33]. Für den Fall, dass sich dieses Produkt durchsetzt, könnte es aufgrund seines Konzeptes auch für die Forschung im Living Place interessant sein.



Abbildung 7: Gatebox

3 Designüberlegungen

Diese Arbeit legt den Fokus auf die technische Umsetzung eines sozialen Agenten. Technik ist jedoch immer nur Mittel zum Zweck, welcher in diesem Fall die Interaktion und der Beziehungsaufbau mit dem Menschen ist. Daher ist es unumgänglich zuerst den Menschen selbst zu betrachten und nach einem kurzen Exkurs in die Produktentwicklung ausgewählte Felder der Psychologie zu beleuchten. In Bezug auf soziale Agenten wird dazu zuerst ein allgemeines Designvorgehen aus der Produktentwicklung vorgestellt. Anschließend folgen Betrachtungen der grundlegendsten Fähigkeiten und Eigenschaften eines sozialen Agenten aus Sicht der Psychologie. Im Bereich der Fähigkeiten werden die Grundlagen der Kommunikation und des Beziehungsaufbaus vorgestellt. Anschließend wird im Bereich der Eigenschaften auf den Charakter sowie das Erscheinungsbild des Agenten eingegangen.

3.1 Allgemeines Designvorgehen

Der Gestaltung sozialer Agenten sind keine Grenzen gesetzt. Der Agent kann durchaus Eigenschaften und Fähigkeiten besitzen, die weit entfernt von bisher bekannten Interaktionspartnern sind. Wichtig bei der Interaktion von Menschen und Agent ist einzig, dass das Konzept des Agenten in sich stimmig ist. Dabei ist es von Vorteil, jedoch nicht zwingend notwendig, dass die Interaktion mit dem Agenten intuitiv gestaltet ist [56]. Ob ein Konzept stimmig ist, kann letzten Endes nur durch Praxistests in Erfahrung gebracht werden. Ein etabliertes Designvorgehen mit integrierten Tests ist das User-Centered Design [12]. Es stellt die User Experience des Produktes in den Vordergrund, welche die gesamte Erfahrung des Nutzers vor, während und nach der Interaktion mit dem Agenten umfasst. In seinen Grundzügen besteht es aus drei Aspekten; dem frühen und beständigen Fokus auf die Nutzer, der frühen und empirischen Messung von Nutzungsdaten und dem iterativen Designprozess.

3.2 Kommunikationsgrundlagen

Die Interaktion mit dem Menschen ist die fundamentale Fähigkeit des sozialen Agenten. Da sich in dieser Arbeit auf die Kommunikation als einzige Interaktionsform beschränkt wird, ist es essenziell deren Grundlagen zu kennen. Zuerst soll die Frage beantwortet

werden, was und wie ein Mensch kommuniziert. Dazu werden im Folgenden zwei Kommunikationsmodelle und deren Aussagen für die Gestaltung sozialer Agenten vorgestellt. Anschließend wird auf die Grundzüge der menschlichen Emotion eingegangen, da diese bei der Kommunikation eine wesentliche Rolle spielen [70].

Kommunikationspyramide nach Mehrabian

Dieser Abschnitt gibt Antworten auf die Frage, wie der Mensch kommuniziert. Für die Antwort wird die Kommunikationspyramide nach Mehrabian [3] herangezogen. Diese unterscheidet zwischen der verbalen, der paraverbalen und der nonverbalen Kommunikation. Die verbale Kommunikation ist hierbei das gesprochene Wort. Die paraverbale Kommunikation betrachtet, wie etwas ausgesprochen wird, und welche weiteren Laute der Mensch von sich gibt. Jede Kommunikation, die keine akustischen Signale benötigt, wird als nonverbale Kommunikation definiert. Mimik, Gestik und Berührungen sind dieser Kommunikationsart zuzuordnen. Nach Mehrabian besteht die Kommunikation des Menschen zu 7% aus der verbalen Kommunikation, zu 38% aus der paraverbalen Kommunikation und zu 55% aus der nonverbalen Kommunikation. Nicht in diesem Modell erfasst ist die extraverbale Kommunikation. Sie beschreibt sämtliche Rahmenbedingungen der Interaktion. Hierzu zählen Körperhaltung, Kleidungsstil, Abstand der Kommunikatoren voneinander, Uhrzeit und vieles Weitere. Die Wichtigkeit der extraverbale Kommunikation darf nicht unterschätzt werden, da diese vielfach festlegt, ob es überhaupt zu einer Kommunikation kommt [26]. Weiterhin sei erwähnt, dass Kommunikation auch innerhalb einer Kommunikationsart sowohl explizit als auch implizit erfolgen kann [34]. So kann nicht nur ein Lächeln, sondern auch eine erhöhte körperliche Aktivität oder das wiederholte Erwähnen des positiven Ereignisses die Information „Freude“ vermitteln. Für den kommunikationsfähigen sozialen Agenten kann somit geschlussfolgert werden, dass beim Kommunikationsdesign eine Vielzahl von Faktoren beachtet werden sollten, da die verbale Kommunikation nur einen sehr geringen Bereich des gesamten Informationsspektrum erfasst.

Kommunikationsquadrat nach Schulz von Thun

Zur Beantwortung der Frage nach dem, was der Mensch kommuniziert, wird das Kommunikationsquadrat [81] nach Schulz von Thun betrachtet. Nach diesem kommuniziert der Mensch immer auf vier Ebenen; der Sachebene, welche den Gesprächsinhalt umfasst;

der Beziehungsebene, welche den Beziehungsstatus des Sprechers zu dem Gegenüber abbildet; der Apellebene, auf welcher kommuniziert wird, was der Sprecher von dem Gegenüber verlangt; und der Selbstoffenbarungsebene, auf welcher der Sprecher seinen eigenen Status preis gibt. Es ist nicht möglich, auf einer dieser Ebenen nicht zu kommunizieren, weswegen folglich alle vier bei dem Design eines sozialen Agenten beachtet werden müssen.

Grundemotionen

Emotionen sind ein wichtiger Bestandteil der menschlichen Kommunikation [70]. Somit stellt sich die Frage, über welche kommunikativen Fähigkeiten ein sozialer Agent in Bezug auf Emotionen besitzen sollte. Auch wenn es keine offizielle Definition gibt, was eine Emotion ist, so gibt es doch mehrere relevante Forschungsergebnisse zu diesen. Eines der etabliertesten kommt von dem Psychologen Dr. Paul Ekman. Er hat in der Mimik sechs Grundemotionen entdeckt, die sich bei allen Menschen auf der Welt gleichen [27]. Weiterhin hat Robert Plutchik acht Grundemotionen definiert und diese gemäß ihren Abhängigkeiten zueinander kreisförmig angeordnet. Zusätzlich beachtet er ihre Intensität, indem er diesen Kreis mit dem Abschwächen der Emotionen zu einem Punkt hin zusammenlaufen lässt [60]. Beide Forscher überschneiden sich in ihren Aussagen in den Grundemotionen „Freude“, „Angst“, „Trauer“ und „Wut“ [62]. Diese Emotionen können damit als essenziell für die Kommunikation angesehen werden. Somit sollte auch ein sozialer Agent über diese Kommunikationsfähigkeiten verfügen, indem er diese zu deuten und auszustrahlen vermag. Für die Verarbeitung dieser Emotionen kann weiterhin das Modell von Plutchik eine Orientierung bieten.

3.3 Beziehungsaufbau

Ein definiertes Ziel sozialer Agenten ist der Beziehungsaufbau zu einem Menschen. Aus diesem Grund wird im Folgenden näher auf die Grundlagen des Beziehungsaufbaus eingegangen. Bei diesem Ziel ist anzumerken, dass die Definition der sozialen Agenten nicht spezifiziert, wie tief die Beziehung zwischen Menschen und Agenten zu sein hat. Damit bleibt offen, wie ausgeprägt die Beziehungsfähigkeit des sozialen Agenten sein muss. Das Interesse der Menschen an der Kontaktaufnahme mit einer sozialen Maschinen ist aufgrund von Neugier im Allgemeinen gegeben. Somit wird es dem sozialen Agenten leicht gemacht, eine Vielzahl an flachen Beziehungen zu Menschen aufzubauen. Aufwändiger ist

es jedoch, tiefere Beziehungen zwischen Agenten und Menschen zu erreichen. Agenten, welche über diese Fähigkeit verfügen, werden auch Relational Agents genannt [14]. Auf die Grundlagen dieser Relational Agents wird im Folgenden weiter eingegangen. Dazu wird der Begriff der zwischenmenschlichen Beziehung im Generellen betrachtet und dabei auf wichtige Aspekte in Bezug auf soziale Agenten eingegangen. Abschließend wird weiterhin auf besondere Tücken bei der Beziehung von Mensch und Maschine hingewiesen.

Zwischenmenschliche Beziehungen können sehr unterschiedlich ausfallen. Es gibt daher keine wissenschaftlich anerkannten allgemeinen Definitionen, wie eine Beziehung auszusehen hat oder wie sie aufzubauen ist. Es können jedoch einige Faktoren für das Gelingen eines Beziehungsaufbaus bestimmt werden. In Bezug auf Arbeitsbeziehungen spezifischer Berufsbilder, wie beispielsweise Pflegekräfte [59], gibt es Leitlinien, an denen sich orientiert werden kann. Weiterhin existieren mehrere Aspekte, die bei dem Beziehungsaufbau und dessen Erhalt beachtet werden sollten. Ein Grundstein jeder positiven Beziehung ist Vertrauen, dies gilt im Besonderen auch für die Beziehung von Mensch und Maschine [13]. Damit die Beziehung erhalten werden kann, ist weiterhin das Interesse an dem Gegenüber wichtig. Dies ist der Punkt an dem Vector und Jibo gescheitert sind, da sie auf Dauer in ihrem Verhalten zu unattraktiv waren [85]. Für das Wachstum der Beziehung sorgen unter anderem das gegenseitige Verständnis und das gegenseitige Helfen [40]. Das Verständnis des Gegenübers ist besonders auf emotionaler Ebene wichtig, was dem sozialen Agenten die Fähigkeit zur Empathie abverlangt. Eine umfassende Studie zu der Umsetzung der Empathie in Bezug auf soziale Agenten hat die Professorin für Informatik Elisabeth Andre in ihrer Arbeit *„Lässt sich Empathie simulieren? Ansätze zur Erkennung und Generierung empathischer Reaktionen anhand von Computermodellen“* verfasst [5]. Für ausführliche Informationen zu diesem Thema sei auf diese Studie verwiesen. Das gegenseitige Helfen kann darüber hinaus sehr vielfältig ausfallen. Wichtig ist nur, die Möglichkeit auch dem Agenten helfen zu können, bei dem Entwicklungsprozess eines Relational Agents zu berücksichtigen. Eine Übersicht und Empfehlungen für den Entwurf von sozialen Robotern für Langzeitinteraktionen bietet das Paper von Leite, Martinho und Pavia *„Social robots for long-term interaction: a survey“* [45].

Es gibt mindestens zwei Gründe, warum die Beziehung zwischen Mensch und Agenten nicht zu tief werden sollte [40]. Zum einen besteht die Gefahr, dass der Mensch eine zu emotionale Bindung zu dem Agenten entwickelt und somit beginnt über die Tatsache hinwegzusehen, dass es sich bei dem Agenten um eine ersetzbare und gefühlsunfähige Maschine handelt. Dies kann zu folgenschweren Fehlinterpretationen und damit Fehlent-

scheidungen seitens des Menschen führen. Wann diese Schwelle überschritten ist, hängt sehr vom Menschen und dessen sozialem Status ab. Die zweite Motivation für das absichtliche Hemmen der Beziehungsentwicklung zum Menschen ist die soziale Wahrnehmung des Agenten durch den Menschen. Hat sich der Agent erfolgreich in das Sozialgefüge integriert, wird dieser auch als Akteur wahrgenommen. An den Agenten werden sämtliche sozialen Anforderungen eines Menschen gestellt, mit denen dieser nun umzugehen hat. Es ist darum wichtig, klare Linien zu ziehen, wie weit eine Beziehung zu einem Menschen und die Integration in sein soziales Umfeld gehen darf.

3.4 Charakter

Um den Menschen langfristig zu der Interaktion mit einem Roboter zu motivieren, benötigt der Roboter einen Charakter [40]. Welche Eigenschaften der Charakter umfasst und wie dieser an einem sozialen Agenten adaptiert werden kann, erklärt dieses Unterkapitel.

Zunächst ist anzumerken, dass keine einheitliche Definition von Charakter existiert. In dieser Arbeit wird der Charakter mit dem Begriff Persönlichkeit gleichgesetzt. Er wird behelfsweise als Summe wiederkehrender Verhaltenseigenheiten definiert. Diese Verhaltensmuster können nun unterschiedliche Tendenzen haben. Für die Kategorisierung dieser charaktergebenden Tendenzen gibt es bereits mehrere Modelle. In Bezug auf soziale Agenten findet häufig das DISC-Modell [68] oder das Big Five Personality Factors Modell [67] Verwendung. DISC steht dabei für die vier Persönlichkeitsfelder „*Dominance*“, „*Influence*“, „*Steadiness*“ und „*Conscientiousness*“, welche von den zwei Achsen Extrovertiert-Introvertiert und Menschenorientiert-Aufgabenorientiert aufgespannt werden. Bekannt ist dieses Modell auch unter dem Namen Vier-Farben-Modell. Das Big Five Personality Factors Modell baut sich aus den Parametern „*Openness*“, „*Conscientiousness*“, „*Extraversion*“, „*Agreeableness*“ und „*Neuroticism*“ (emotionale Stabilität) auf. Die individuellen Ausprägungen dieser Parameter beschreiben dabei die Persönlichkeit eines Mensch. Aufgrund der Anfangsbuchstaben dieser Parameter wird es auch als OCEAN-Modell bezeichnet.

Die Adaption des Charakters erlaubt es dem Agenten seine Eignung für wechselnde Aufgaben, Situationen und Interaktionen zu erhöhen. Für die Anpassung des Charakters können verschiedenste Verhaltensfaktoren angepasst werden. In der weiteren Ausführung

wird sich dazu auf das Paper „*User-robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy*“ [80] von Tapus, Tapus und Mataric bezogen. In diesem hat sich beispielsweise gezeigt, dass die Bereitschaft steigt mit einem verspielten Agenten zu interagieren, während mit ernsthaften Agenten die Bereitschaft zur Kooperation steigt. Jedoch kann das Verhalten nicht nur an die Situation, sondern auch an den Interaktionspartner angepasst werden. Generell gilt, dass der Mensch dem sozialen Agenten am meisten zugeneigt ist, der seinem eigenen Charakter am meisten entspricht. Dazu können verhaltensspezifische Faktoren entsprechend adaptiert werden. Auf verbaler und paraverbaler Ebene kann der Charakter durch den Inhalt, die Betonung und die Sprachgeschwindigkeit variiert werden. Auf nonverbaler Ebene kann dieser durch den Grad der Aktivität des Körpers und den Abstand zum Nutzer angepasst werden. Wie sehr der Charakter des Agenten dem Nutzer entspricht, kann unter anderem durch die Interaktionsdauer und die Anzahl freiwilliger Interaktionen festgestellt werden.

3.5 Äußere Erscheinung

Die äußere Erscheinung des sozialen Agenten deckt einen großen Bereich der extraverbaler Kommunikation ab [26]. Sie ist entscheidend dafür, wie der soziale Agent vom Menschen eingeordnet wird und in welchem Maß der Mensch zu der Interaktion mit der Maschine bereit ist. Dieses Unterkapitel legt daher die wichtigsten Phänomene sozialer Agenten in Bezug auf ihr Erscheinungsbild dar. Die äußere Erscheinung eines sozialen Agenten umfasst dabei sämtliche Aspekte des oberflächlichen Eindrucks auf den Menschen und beschränkt sich nicht allein auf das Optische. Begonnen wird mit dem Einfluss der Körperhaftigkeit. Fortgefahren wird mit den starken Einflüssen des Antromorphismus und des Kindchenschemas. Abgeschlossen wird mit dem Phänomen des Uncanny Valleys.

Körperhaftigkeit

Was die Körperlichkeit des sozialen Agenten angeht, sei noch einmal betont, dass der erfolgreiche soziale Agent keinen Körper besitzen muss. Es hat sich jedoch gezeigt, dass sozialen Agenten bei der Interaktion mit Menschen eine höhere Relevanz zugeschrieben wird, wenn diese über einen Körper verfügen. Die Qualität der Interaktion bei körperhaften Agenten ist tendenziell höher als bei körperlosen Agenten [56]. Dieser Effekt wird verstärkt, wenn der Körper auch Bewegungen ausführen kann [61]. Der Nachteil an einem

Körper ist jedoch der erhöhte Umsetzungsaufwand und die Möglichkeit zur Kollision, sollte dieser ebenfalls bewegungsfähig sein. Ein guter Kompromiss für die Umsetzung eines bewegungsfähigen Körpers kann daher die Darstellung auf einem Bildschirm sein. Die zugesprochene Relevanz ist in diesem Fall geringer als bei einem realen Körper, dafür steigt jedoch die Einsatzsicherheit und die Einfachheit der Umsetzung erheblich.

Vermenschlichung

Anthropomorphismus ist der Fachausdruck für die Neigung des Menschen in allen Dingen menschliche Züge wiederzuerkennen. Vermenschlichung wird dieses im Volksmund genannt. Unterbewusst neigen Menschen bei sämtlichen Interaktionen dazu, den Interaktionspartner als menschlich zu deuten. Soziale Agenten können sehr von dieser tief im Menschen verankerten Eigenschaft profitieren [31]. Es reichen damit schon vage Andeutungen menschlicher Eigenschaften, um den Agenten als sozialen Interaktionspartner in Szene zu setzen. Der Anthropomorphismus kann jedoch auch zum Problem werden, wenn der Agent seinem äußerlichen Eindruck nicht entsprechen kann [85]. Erkennt der Mensch beispielsweise ein Gesicht an dem sozialen Agenten, so könnte jener die Erwartung entwickeln, dass der Agent sein Gesicht dem Menschen bei der Interaktion zuwendet. Kann der soziale Agent diese Erwartung nicht erfüllen, wird der Mensch im wahrsten Sinne des Wortes enttäuscht. Er erhält den Eindruck, dass etwas mit seinem Gegenüber nicht stimmt. Dies kann zu Desinteresse an der weiteren Interaktion oder sogar Abneigung dem Agenten gegenüber führen. Aus diesem Grund ist es wichtig, die äußere Erscheinung des Agenten mit viel Bedacht zu entwerfen.

Kindchenschema

Eine bewährte Möglichkeit, die Ansprüche des Nutzers an die Fähigkeiten des sozialen Agenten zu senken, bietet das Kindchenschema [44]. Erscheint der Agent kindlich, erhöht sich die Bereitschaft des Menschen Fehler zu verzeihen. Gleichzeitig erzeugt das Kindchenschema bei dem Menschen das Bedürfnis, sich um den Agenten kümmern zu wollen. Das Interesse an der Interaktion und die Zuneigung zu dem Agenten wird damit gesteigert. Dieser Effekt kann durch soziale Agenten so erfolgreich eingesetzt werden, dass es bereits erste Bedenken zu der ethischen Richtigkeit dieser Täuschungstechnik gibt [44]. Nachteilig an dem Kindchenschema ist jedoch, dass es sich per Definition nur durch die Körperproportionen auszeichnet. Besitzt der Agent keinen Körper, ist zu prüfen in wie

weit ähnliche Effekte über andere kindliche Merkmale, wie beispielsweise die Stimmlage, erzeugt werden können.

Uncanny Valley

Kann der Mensch objektiv nicht mehr unterscheiden, ob es sich bei dem Agenten um einen Menschen oder eine Maschine handelt, wird der Mensch versuchen den Agenten menschlich einzuordnen. Problematisch ist dies, da Menschen hierbei sehr feinfühlig vorgehen. Bei der kleinsten Abweichung in Erscheinung oder Verhalten erhält der Mensch den Eindruck, dass mit seinem Gegenüber etwas nicht stimmt. Da der Mensch objektiv jedoch nicht zu dem Schluss kommen kann, dass es sich bei seinem Gegenüber um eine Maschine handelt, übernehmen seine Instinkte die Interpretation. Der Mensch reagiert mit Misstrauen, Furcht, Abneigung und sogar Ekel dem Agenten gegenüber. Dieses Phänomen wurde von Masahiro Mori, einem Japanischen Forscher im Bereich humanoider Roboter, als Uncanny Valley [53] bezeichnet. Selbiges Phänomen kann auch bei Wachsfiguren oder Charakteren in Animationsfilmen auftreten. Mit dem Überschreiten der Schwelle des fehlerhaften Lebewesens, schwächt sich dieses Phänomen wieder ab und die künstlichen Wesen werden vom Menschen akzeptiert. Bei dem Entwurf sozialer Agenten ist es damit wichtig zu beachten, auf welcher Seite des Uncanny Valleys der Agent stehen soll.

4 Kernfunktionen sozialer Agenten

Die technische Umsetzung der Kernfunktionen eines sozialen Agenten sei in diesen Kapitel vorgestellt. Dazu muss zuerst die Frage beantwortet werden, was genau die Kernfunktionen eines sozialen Agenten im Rahmen dieser Arbeit sind. Dies geschieht, indem die Aussagen über die Kernfunktionen eines allgemeinen Agenten mit denen über die Kernfähigkeiten eines sozialen Agenten in Verbindung gebracht werden. Die Kernfunktionen eines allgemeinen Agenten bestehen aus dem Wahrnehmen seiner Umgebung und aus dem Handeln in dieser. Die Kernfähigkeit des sozialen Agenten ist die soziale Interaktion mit dem Menschen, wobei sich in dieser Arbeit auf die Kommunikation mit dem Menschen beschränkt wird. Damit ergibt sich, dass die Kernfunktionen eines sozialen Agenten das Wahrnehmen und Handeln in Bezug auf die Kommunikation mit dem Menschen sind. Dies lässt sich bei Betrachtung der Kommunikationskanäle des Menschen weiter konkretisieren. Die Kommunikationskanäle eines Menschen begründen sich aus drei Ebenen; die Akustik, die Optik und die Berührung. Die ersten beiden Ebenen lassen sich sauber in Wahrnehmung und Handlung auftrennen, während dies bei der Berührung nicht der Fall ist. Es ergeben sich damit fünf maßgebliche Funktionen für die Ausfüllung der Kernfähigkeit des sozialen Agenten. Dieses Kapitel stellt diese in der Reihenfolge akustische Wahrnehmung, optische Wahrnehmung, akustische Handlung, optische Handlung und Berührungen vor.

4.1 Akustische Wahrnehmung

Über die akustische Wahrnehmung des Menschen können eine Vielzahl von Informationen über diesen gewonnen werden. Neben den verbalen Informationen gibt der Sprecher auch paraverbale und sogar augenscheinlich extraverbale Informationen über sich preis. Dieses Unterkapitel bietet eine Einführung in die Möglichkeiten der akustischen Wahrnehmung eines sozialen Agenten. Zuerst wird der Bereich des Sprachverständnisses in seinen Grundzügen erläutert. Darauf folgt eine Übersicht über das Gebiet der Emotionserkennung. Geendet wird mit Einblicken in die Sprecheridentifikation und die Sprecherlokalisierung.

4.1.1 Sprachverständnis

Die wohl naheliegendste Funktion der akustischen Wahrnehmung des Menschen ist das Verstehen dessen Sprache. Dazu müssen die aufgenommenen Sprachlaute in die kommunizierten Informationen umgewandelt werden. Der genaue Ablauf lässt sich in zwei Stufen unterteilen. Zuerst werden die Verfahren der Automatic Speech Recognition (ASR) [91] angewandt, um die Transformation Speech-to-Text zu realisieren. Anschließend werden Satztabellen oder das sogenannte Natural Language Understanding [21] bemüht, welche die Informationen aus diesen Texten extrahieren. Beide Stufen werden im Folgenden näher erläutert. Daraufaufgehend wird auf Probleme und Vereinfachungen zu diesen Verfahren eingegangen.

Die in Stufe eins angewandte Automatic Speech Recognition erfolgt in mehreren Schritten. Ist die Lautabfolge aufgenommen, wird diese in ihre Frequenzen zerlegt. Daraus resultierende Frequenzabfolgen werden mittels eines sogenannten Acoustic Models den Phonen, sprich den kleinsten Bestandteilen einer Sprache, zugeordnet. Anschließend werden diesen Phonen mittels eines Language Models Buchstaben zugewiesen, aus welchen sich wiederum die Wörter und Sätze ergeben. Anschließend werden die Worte und Sätze auf ihre formelle Sinnhaftigkeit und Korrektheit überprüft. Wird die Prüfung nicht bestanden, kommt das Language Modell erneut zum Einsatz und ordnet den Phonen die nächstwahrscheinlichen Buchstaben zu, bis Sinnhaftigkeit und Korrektheit erreicht sind [91].

Die Informationsgewinnung in Stufe zwei kann entweder statisch oder dynamisch erfolgen. Der statische Ansatz beinhaltet eine Lookup-Tabelle, in welcher die Wörter und Sätze zusammen mit den entsprechenden Informationen abgespeichert sind. So einfach dieses Modell in seinem Konzept scheint, so begrenzt ist es jedoch auch in seiner Flexibilität. Diese Begrenzung aufzuheben ist die Motivation für den dynamischen Ansatz. Natural Language Understanding nutzt wesentlich komplexere Verfahren, um die Sätze syntaktisch und semantisch zu analysieren und dadurch ohne Vorwissen die Informationen aus sämtlichen Texten extrahieren zu können. Natural Language Understanding ist dabei ein Teil des Bereiches Natural Language Processing (NLP) [21].

Aktuelle Probleme beim Verstehen von Sprache sind vor allem der Einfluss von Störgeräuschen bei der Sprachaufnahme und der Umgang mit Dialekten und Umgangssprache [91]. Die Anwendungsmöglichkeiten der Informationsübertragung mittels der Sprache sind damit zurzeit noch eingeschränkt.

Um diese Probleme zu umgehen, kann, je nach Konzept des Agenten, die Funktion der Spracherkennung sehr vereinfacht werden. Sowohl die Fehleranfälligkeit als auch der Umsetzungsaufwand sinkt erheblich, wenn die Laute, auf die eine Reaktion des Agenten gewünscht ist, stark eingeschränkt werden. Das Spektrum der Laute ließe sich beispielsweise auf das der effektiven Kommunikation von Mensch zu Hund limitieren. Damit wird erreicht, dass die Laute in Anzahl gering und im Klang eindeutig sind. Die Notwendigkeit für ein komplexes Sprachverständnissystem fällt damit weg.

Weiterhin ist der Aufbau des Sprachverständnisses in Etappen möglich. Aus dem konventionellen Ansatz lässt sich deuten, dass sich hinter jedem VoiceBot ein Chatbot verbirgt. Die Funktion des auditiven Verstehens wird dem Funktionsbereich des Chatbots lediglich hinzugefügt. End-to-End Ansätze sind heutzutage jedoch weit darin fortgeschritten, den konventionellen Ansatz zu ergänzen oder auch zu ersetzen. Die Möglichkeit, die Entwicklung auf diese Weise zu etappieren, muss somit von Fall zu Fall geprüft werden.

4.1.2 Emotionserkennung

Wie erläutert, stellt die Sachebene nur einen Teil der preisgegebenen Informationen in einer menschlichen Kommunikation dar. Es ist daher sinnvoll, nicht nur die verbale, sondern auch die paraverbale Kommunikation zu beachten. Das Verstehen des Menschen auf paraverbaler Ebene wird durch die Emotionserkennung möglich. Hierzu bietet der Themenbereich der Speech Emotion Recognition oder Speech Emotion Detection diverse Möglichkeiten. Eine Übersicht des Themenbereichs kann in dem Survey-Paper von Schuller „*Speech Emotion Recognition: Two Decades in a Nutshell, Benchmarks, and Ongoing Trends*“ [71] gefunden werden. Ausführliche Informationen bietet weiterhin das Buch „*Robust Emotion Recognition using Spectral and Prosodic Features*“ von Rao und Koolagudi [63]. Zu der Umsetzung der Speech Emotion Recognition bedient man sich im allgemeinen Machine Learning Verfahren. Problematisch bei der Emotionserkennung von Sprache über Machine Learning Verfahren sind jedoch die Trainingsdaten [87]. Menschen fällt es schwer, Emotionen aus einem gehörten Text eindeutig zu bestimmen. Sie können sie die Datensätze daher nur eingeschränkt korrekt labeln. Die Alternative ist, die Sprachsequenzen mit den gewünschten Emotionen selbst zu erzeugen. Dazu muss jedoch erheblicher Aufwand betrieben werden, da bei vorgetäuschten Emotionen meist übertrieben wird. Folglich lernt der Algorithmus nicht, feinere und viel häufiger auftretende Emotionsregungen wahrzunehmen. Gilt es daher Testdaten zu generieren, sollten

Menschen mit Bildern, Videos oder Computerspielen zu aufrichtig emotional besetzten Aussagen geführt werden.

4.1.3 Sprecheridentifikation

Besonders wenn der Agent eine tiefere Beziehung zu mehreren Menschen aufbauen soll, kann eine Unterscheidung der Einzelperson durch Identifikation sinnvoll sein. Dies kann über die Bestimmung des Alters und des Geschlechts einer Person geschehen, exakter ist jedoch die Identifikation über die individuellen Besonderheiten einer Stimme. Bedingt sind diese Besonderheiten durch die Einzigartigkeit der biometrischen Eigenschaften des Vokaltrakts eines jeden Menschen. Der Bereich Speaker Recognition [58] konzentriert sich auf die Extraktion und Interpretation dieser Merkmale.

4.1.4 Sprecherlokalisation

Eine weitere Information, die aus den akustischen Signalen eines Menschen ermittelt werden kann, ist dessen Position zum Agenten. Besonders relevant ist dies, sobald der Agent über die Möglichkeit verfügt sich nach dem Menschen hin auszurichten. Der Agent kann so seine Mikrophone optimal zu dem Menschen hin positionieren, um eine bessere Aufnahmequalität zu erreichen. Aber auch aus sozialpsychologischer Sicht ist das Ausrichten zum Gesprächspartner wichtig. Mehr dazu wird in Unterabschnitt 4.4 erklärt. Der etablierte Funktionsbegriff für die Umsetzung dieser Fähigkeit ist die Sound Localization. Eine wichtige Voraussetzung für das Orten von Geräuschen ist, dass der Agent über mindestens zwei Mikrophone in unterschiedlicher Ausrichtung verfügt. Einen Überblick über verschiedene Methoden und die Darstellung eines Umsetzungsbeispiels können in dem Paper von Mahadev und Austin mit dem Titel „*Sound localization by robot using inter-aural time differences*“ [50] gefunden werden. Zu beachten ist bei diesem Umsetzungsbeispiel, dass lediglich die Richtung, nicht aber die Entfernung des Sprechers ermittelt wird. Die Entfernungserkennung über Mikrophone ist jedoch ebenfalls möglich.

4.2 Optische Wahrnehmung

In der Kommunikation spielen optische Signale eine große Rolle. Nahezu die gesamte non-verbale Kommunikation, aber auch große Teile der extraverbalen Kommunikation, erfolgen über den Austausch optischer Signale. Die Wahrnehmung optischer Signale ist damit

für die Konzeptionierung sozialer Agenten eine nicht zu vernachlässigende Fähigkeit. Dieses Unterkapitel bietet eine Einführung in die gängigsten Wahrnehmungsfähigkeiten eines sozialen Agenten. Zuerst wird auf die Positionserkennung und die Personenidentifikation eingegangen. Anschließend wird der große Bereich der Wahrnehmung der Körpersprache vorgestellt. Dies ist hierzu weiter untergliedert in die Aspekte Mimik, Gestik und Körperstellung.

4.2.1 Positionserkennung

Eine grundlegende Funktion innerhalb der optischen Wahrnehmung ist die Positionserkennung. Für die meisten Bildanalyseprozesse ist es zuerst vonnöten, die relevanten Bereiche auf dem Bild zu erkennen, bevor weitere Analyseprozesse ausgeführt werden können. Diese Erkennung erfolgt über das Abgleichen charakteristischer Merkmale des Objektes oder des Menschen mit dem Bild [78]. Weiterführend kann die Position im Bild auf die Position im Raum übertragen werden. Diese Lokalisierung wird über ein mathematisches Modell mit Daten über Entfernung oder Größe einer Silhouette ausgehend von einem Referenzpunkt durchgeführt. Abgesehen von einer Kamera eignen sich für die Erhebung dieser Daten auch ein sogenannter LiDAR oder eine Tiefenkamera. Ein Umsetzungsbeispiel auf Basis einer Tiefenkamera ist in dem Paper „*Mobile Human-Robot Teaming with Environmental Tolerance*“ [47] von Loper, Koenig, Chernova, Jones und Jenkins zu finden.

4.2.2 Personenidentifikation

Neben der Erkennung von Mensch und Objekt liegt die Fähigkeit auch Individuen identifizieren zu können nahe. Es gibt viele optische Merkmale, an Hand derer ein Mensch identifiziert werden kann. Wie in der akustischen Wahrnehmung können auch hier extraverbale Daten wie Alter oder Geschlecht erhoben werden, um eine Person zu identifizieren. Es gibt jedoch wesentlich aussagekräftigere Ansätze. Ein prominentes Beispiel ist dabei die Gesichtserkennung. Umgesetzt wird diese allgemein in drei Schritten [32]. Zuerst wird das Gesicht auf dem Bild lokalisiert, dieser Schritt wird auch Facial Detection genannt. Anschließend werden in diesem Gesicht biometrisch relevante Punkte markiert und ihre Relationen zueinander ermittelt. Diese Relationen werden im dritten Schritt, wie ein Hashwert, mittels einer Lookup-Tabelle einer einzelnen Person zugeordnet. Die

letzten beiden Schritte werden auch zum Facial Recognition zusammengefasst. Für ausführlichere Informationen sei das Paper „*A Survey On Face Recognition Algorithms*“ [4] verwiesen.

4.2.3 Mimik

Über die Mimik kann der soziale Agent mindestens zwei Informationen deuten. Zum einen sind dies die Emotionen und zum anderen die Aufmerksamkeit des Menschen. Die Emotionserkennung aus den Gesichtszügen bietet die beste Auskunft über den emotionalen Zustand eines Menschen [27]. Die Facial Emotion Recognition bildet einen Teilbereich des Themengebietes der Facial Expression Recognition [22]. Ein konventionelles Verfahren der Facial Expression Recognition läuft in seinen Grundzügen wie folgt ab. Nachdem das Gesicht lokalisiert ist, werden einzelne Bereiche eines Gesichtes identifiziert und deren Abstand zueinander bestimmt. Von den Abständen der Gesichtspartien zueinander wird auf die Anspannung verschiedenster Gesichtsmuskeln zurückgeschlossen. Die identifizierten angespannten Muskeln können dann mit einer Datenbank abgeglichen werden, welche die Kombination und den Grad der angespannten Gesichtsmuskeln den kommunizierten Informationen des Menschen gegenüberstellt. Alternativ zu diesem Verfahren existieren verschiedenste End-to-End Ansätze. Die End-to-End Ansätze besitzen jedoch im Vergleich zu konventionellen Ansätzen den Nachteil, dass sie die Gesichtszüge nur den im Training vorgegebenen Emotionen zuordnen können. Eine Mischung dieser Emotionen kann damit mitunter nicht erkannt werden. Weiterhin ist bei einem End-to-End Ansatz die Bereitstellung der Trainingsdaten problematisch. Das Labeln von Bildern fällt dem Menschen wesentlich einfacher als bei der akustischen Emotionserkennung, jedoch bleibt die Schwierigkeit der Authentizität der Emotionen beim Erzeugen neuer Trainingsdaten bestehen. Die zweite Information, die der soziale Agent über die Mimik erkennen kann, ist die Aufmerksamkeit des Menschen. Die Aufmerksamkeitserkennung erfolgt dabei über das Analysieren der Ausrichtung der Augen. Das was der Mensch anblickt, darauf richtet sich seine Aufmerksamkeit. Diese Analyse ist die Kernfunktion des Bereichs Eye Tracking [51].

4.2.4 Gestik

Die Gestik ist definiert als die konversationsbezogenen Bewegungen des Oberkörpers eines Menschen. Diese automatisiert zu interpretieren ist die Aufgabe der Gesture Recognition

[89]. Viele Ansätze sind dabei auf die Hände des Menschen fokussiert. Dies ist damit zu erklären, dass die Hände die komplexesten Gesten ausführen können und ihre Position gleichzeitig Aufschluss über die Bewegung des restlichen Oberkörpers zulassen. Hände bieten weiterhin den Vorteil, dass sie sich farblich meist aus dem Bild hervorheben. Beispielsweise wird zuerst das Verfahren des Face Detections verwendet, um aus dem Gesicht den Farbton der Haut des Menschen zu ermitteln. Anschließend werden alle Bereiche mit identischen Farben isoliert und das Gesicht extrahiert. In den übrigbleibenden Regionen werden nun die Hände lokalisiert. Auf Grundlage dieser Daten erfolgt nun die Gestenerkennung. Hierzu werden biometrische Merkmale einer Hand dem Bild zugeordnet und diese in Relation zueinander gesetzt. Die Geste lässt sich dann diesen Relationswerten zuordnen und kann somit erkannt werden. Eine Übersicht über den aktuellen Stand der Technik kann in dem Paper „*Hand Gesture Recognition: A Survey*“ [90] von Yashas und Shivakumar gefunden werden.

4.2.5 Körperstellung

Besonders für CareBots und Cobots kann das Erkennen der Körperstellung des Interaktionspartners von hoher Relevanz sein. Aber auch allgemein ist das Erkennen der Körperstellung für soziale Agenten von Interesse. Sie erlaubt es, Kontextinformationen über das weitere Umfeld des Agenten zu sammeln oder auch eine Redundanz in der Wahrnehmung zu schaffen. Beispielsweise kann die Körperstellung Aufschluss über den emotionalen Zustand des Menschen geben [77]. Der Themenbereich der Pose Estimation oder auch Posture Recognition beinhaltet hierbei die Funktionsumsetzung dieser Fähigkeiten [54]. Das Konzept ist ähnlich wie bei der Gesture Recognition. Zuerst werden die relevanten Bereiche in dem Bild lokalisiert. Auf diesen Bereichen werden sodann relevante Punkte eines menschlichen Körpers markiert und aus den Relationen dieser die Körperstellung des Menschen erschlossen. Weiterführend existiert der Bereich der Action Recognition [24], um nicht nur den Menschen, sondern auch dessen Handlung zu erkennen.

4.3 Akustische Handlung

Neben der Wahrnehmung muss ein Agent die Fähigkeit zum Handeln besitzen. Eine Handlungsfähigkeit ist dabei das Aussenden akustischer Signale. Akustische Signale erlauben es dem Agenten, sich sowohl auf sachlicher als auch auf emotionaler Ebene dem

Menschen verständlich zu machen. Damit bilden akustische Signale das grundlegende Element für die Handlungsfähigkeit in der menschlichen Interaktion. Was die Synthese akustischer Signale angeht, so steht dessen Komplexität in Relation zu der Variabilität der auszugebenden Laute. Diese Arbeit unterscheidet daher in die vordefinierte, beziehungsweise statische, Lautsynthese und die nicht vordefinierte, beziehungsweise dynamische, Lautsynthese. Zuerst werden die Möglichkeiten zur statischen Lautsynthese vorgestellt. Danach wird die komplexere dynamische Lautsynthese erläutert und als Erweiterung dieser auf die emotionsbehaftete Lautsynthese eingegangen. Zuletzt werden weiterhin wichtige Aspekte des Kommunikationsbeginns beschrieben.

4.3.1 Statische Lautsynthese

Eine feste Anzahl an vordefinierten Lauten, Wörtern oder Sätzen können ohne großen Aufwand erzeugt werden. Als simpelste Implementation ist beispielsweise eine Lookup-Tabelle denkbar, die einen erreichten inneren Zustand des Agenten einem Laut gleichsetzt. Eine Steigerung in der Variabilität kann weiterhin durch die Kombination vordefinierter Laute erzeugt werden. Eine Umsetzungsmöglichkeit wäre hierbei eine mehrdimensionale Matrix, welche beim Erreichen mehrerer interner Zustände jedem einen Laut gleichsetzt. Die Laute werden hintereinander abgespielt, um so eine Varianz der einzelnen Laute zu erzeugen. Ein Beispiel für diese Anwendung ist die Ansage der Uhrzeit. Abhängig von der Anzahl der Laute und derer Kombinationsmöglichkeiten kann die Komplexität dieser Umsetzung deutlich steigen. Der ansteigenden Komplexität könnte mit der Nutzung von Textgerüsten entgegengewirkt werden. Diese sind vergleichbar mit Lückentexten, welche zu ihrer Vollständigkeit lediglich mit weiteren Lauten, Wörtern oder Satzteilen gefüllt werden müssen. Selbige Verfahren sind auch für die Generierung von Texten einsetzbar.

4.3.2 Dynamische Lautsynthese

Will man sich von den statischen Ansätzen lösen, kommen Methoden zum Einsatz, welche die Lautausgabe völlig frei gestaltbar machen. Verwendung findet dies vor allem in der Synthese der verbalen menschlichen Kommunikation. Um diese Variabilität zu erreichen, wird im ersten Schritt ein auszusprechender Text generiert und dieser im zweiten Schritt in Sprache übertragen. Die dynamische, variantenreiche Texterstellung wird dabei durch künstliche Intelligenzen, den sogenannten Conversational AIs [43], ermöglicht. Der zweite

Schritt besteht aus der Text-to-Speech Transformation. Diese kann heutzutage ebenfalls von den Conversational AIs abgedeckt werden. Alternativ sind in den konventionellen Ansätzen mehrere Funktionen erforderlich. Zuerst wird der Text normalisiert. Das bedeutet, dass sämtliche Symbole wie Nummern, Abkürzungen durch ihre ausgeschriebene Variante ersetzt werden. Anschließend wird der normalisierte Text in Teile zerlegt und als aneinandergereihte Laute ausgegeben. Dieses Aufgabenfeld deckt der Bereich der Speech Synthesis [79] ab. Beide Schritte zählen zu dem Gebiet der Natural Language Generation, welcher wiederum ein Teilbereich des Natural Language Processing [21] bildet.

4.3.3 Emotionsbehaftete Lautsynthese

Die Ebene der paraverbalen Kommunikation kann auf akustischer Ebene erschlossen werden, indem neben der Lautabfolge die mitschwingenden Emotionen variabel gestaltet werden. Auch diesen Bereich deckt die Natural Language Generation ab. Für jede kategorische Emotion muss es dazu eine Variante des Aussprachemodells geben. Weiterführend lassen sich diese Aussprachemodelle für die Verarbeitung von Emotionswerten mit variabler Intensität aufbauen. Dieses Modell ermöglicht nicht nur die Variabilität in der Intensität der Einzelemotionen, sondern erlaubt auch eine Mischung verschiedener Emotionen zu unterschiedlichen Anteilen. Für weitere Informationen wird auf die Arbeit von Burkhardt und Campbell „*Emotional speech synthesis*“ [18] in dem Buch „*The oxford handbook of affective computing*“ verwiesen.

4.3.4 Der Kommunikationsbeginn

Abschließend sei auf die Herausforderungen des Kommunikationsbeginns und dessen Lösungsansätze aufmerksam gemacht. Handelt es sich bei dem sozialen Agenten um ein Always-On System, so muss der Agent selbstständig ermitteln können, wann ein guter Augenblick für einen Gesprächsanfang ist und wann nicht. Dazu muss der Agent das Umfeld und die Handlung des Menschen, mit dem eine Interaktion geplant ist, interpretieren können. Ein Gesprächsbeginn kann aus vielerlei Hinsicht unpassend sein, sodass die Anforderungen an die Wahrnehmungsfähigkeit des Agenten für die Erkennung eines passenden Momentes hoch sein müssen. Eine simple Lösung könnte der Ansatz bieten, bei Kommunikationsbedarf lediglich diskret auf sich aufmerksam zu machen. Wird eine positive Reaktion des Menschen erkannt, kann die Interaktion begonnen werden. Eine weitere Lösung ist es, dass von dem Agenten aus keine Interaktion begonnen wird. Der

Agent wartet in diesem Fall bis der Mensch die Interaktion von sich aus beginnt. Dazu muss der soziale Agent jedoch den vom Menschen ausgehenden Interaktionsbeginn wahrnehmen können. Das Erkennen eines Schlüssellautes oder einer Schlüsselgeste sind hierbei denkbare Verfahren.

4.4 Optische Handlung

Die Definition eines Agenten lässt die Existenz und die Gestaltung seines Körpers offen. Daher stehen diesem vielfältige Möglichkeiten zur Verfügung, einem Menschen optische Signale zu senden. Farben, Formen und Bilder seien hier als alternative nonverbale Kommunikationswege genannt. Die Konzepte zur Umsetzung dieser Kommunikationsformen sind mannigfaltig. Besonders leicht lassen sich gesendete Signale jedoch verstehen, wenn diese dem Menschen nachempfunden sind. Dies macht sie für jeden Menschen intuitiv verständlich. In diesem Kapitel sei daher auf das Nachempfinden der menschlichen optischen Kommunikation eingegangen. Wie bereits dargelegt, kann diese in drei Aspekte unterteilt werden. Begonnen wird mit der Synthese der Mimik, gefolgt von der Synthese der Gestik und der Körperstellung.

4.4.1 Mimik

Die Mimik setzt ein bewegungsfähiges Gesicht voraus. Die einfachste Umsetzungsweise ist es dabei, Gesichtszüge auf einem Display darzustellen. Da der Mensch dazu neigt, in verschiedenste Formen Gesichter hineinzuzinterpretieren, kann das Gesicht des Agenten auch in abstrakten Andeutungen bereits als solches erkannt werden [31]. Die über die Mimik kommunizierten Informationen werden in dieser Arbeit in zwei Aspekte eingeteilt; den emotionalen Zustand und die Aufmerksamkeit. Im Folgenden wird zuerst auf die visuelle Kommunikation der Emotionen und darauffolgend auf die der Aufmerksamkeit eingegangen.

Was die Emotionen angeht, so sind beim Menschen Gesichtsausdruck und Betonung der Sprache miteinander gekoppelt. Verfügt ein sozialer Agent über die Fähigkeit optische und akustische Signale auszusenden, ist darauf zu achten, dass die Emotionen innerhalb dieser Signale zueinander passen [5]. Ausgedrückt werden die Emotionen durch die Form und Position der Gesichtskonturen, wobei besonders Mund und Augen ausschlaggebend sind [10]. Besonders bei der Mimik ist es wichtig, auch einen neutralen Gesichtsausdruck für den Agenten zu kreieren.

Im Bereich der Aufmerksamkeit spielt vor allem die Ausrichtung der Augen eine wichtige Rolle. Nach der Arbeit „*Realistic and Interactive Robot Gaze*“ [57] erhöht ein sozialer Agent die Illusion des interessierten Gesprächspartners, wenn er den Menschen mit den Augen ansieht. Weiterhin kann durch die Ausrichtung der Augen die Handlungsabsicht des Agenten kommuniziert werden. Besitzt der Agent die Möglichkeit der Objektmanipulation oder der Fortbewegung, ermöglicht die Ausrichtung der Augen beispielsweise eine sehr intuitive Aussage über das nächste Aktionsziel des Agenten [16]. Sind in den Augen selbst keine Kameras verbaut, wird für diese Funktion ein mathematisches Modell benötigt, welches aus den Ausrichtungsdaten des Agenten und den Positionsdaten des Menschen die Ausrichtung der Augen ermittelt.

4.4.2 Gestik

Die Gestik des sozialen Agenten kann ebenfalls sehr vielfältig ausfallen. Da ein sozialer Agent nicht zwingend über einen Oberkörper verfügt, wird der Begriff der Gestik hier verallgemeinert auf Bewegungen des Körpers, die nicht einem Gesicht zuordenbar sind, jedoch im unmittelbaren Blickfeld des Menschen liegen. Neben den menschlichen Gesten wird dieser Begriff für soziale Agenten somit auf einem Display, Farben oder auch allgemeine Körperbewegungen erweitert. Beispiel hierfür ist das Display von Pepper [61] oder die farbige Umsetzung der Gestik eines sozialen Roboters in der Arbeit „*Robot-specific social cues in emotional body language*“ [28] von Embgen, Luber, Becker-Asano, Ragni, Evers und Arras. Je nach Größe und Gestalt des Agenten inkludiert die Gestik bereits die gesamten Körperbewegungen, womit die Kategorie der Körperstellung wegfällt; siehe Jibo oder Vector. Sollte der soziale Agent Gestiken mittels physischen Körperbewegungen realisieren, ist es für die Einsatzsicherheit wichtig, das Kollisionsrisiko zu beachten. Ist der soziale Agent ohne feste Absperrungen eingesetzt, muss davon ausgegangen werden, dass eine Kollision mit einem Objekt oder einem Menschen über die Zeit unvermeidlich ist. Die Lösung ist hier die Maximalkraft des Agenten soweit zu beschränken, dass bei der direkten Kollision keine Schäden entstehen [83]. Für weitere Ausführungen zu der Schadensvermeidung bei Kollision sei auf Unterabschnitt 4.5 verwiesen.

4.4.3 Körperstellung

Alle kommunikativen Bewegungen, die nicht durch die Gestik erfasst werden, sind hier zur Körperstellung zusammengefasst. Besonders für Roboter, bei deren Design die soziale

Interaktion nicht im Vordergrund steht, ist die Kommunikation über die Körperstellung interessant [11]. Aber auch soziale Agenten profitieren von der Kommunikation über die Körperstellung. Die Modifikation der Körperstellung erlaubt es ihnen beispielsweise, einem Menschen Bewegungen vorzumachen oder über größere Distanzen mit diesem zu kommunizieren. Auch gestaltet die konsistente Kommunikation auf mehreren Kommunikationswegen den sozialen Agenten lebendiger. So lassen sich beispielsweise Emotionen auch durch die Körperstellung ausdrücken [8]. Kommunizierte Informationen mit Körperbewegungen zu untermalen, wird dabei als proaktive Beweglichkeit bezeichnet. Auch bei dieser Kommunikationsform ist die Kollisionsgefahr zu beachten.

4.5 Berührungen

Bei der Berührung sind Wahrnehmung und Handlung miteinander verbunden. Daher lässt sich dieses Thema am besten vereint betrachten. Welchen Stellenwert die Berührung in der Kommunikation hat, hängt vom Kulturkreis ab. Sie ist jedoch ein wichtiger Bestandteil der engen zwischenmenschlichen Beziehung. Somit können auch soziale Agenten von der Fähigkeit der Berührung profitieren [30]. Weiterhin werden viele Dienste von Servicerobotern, wie beispielsweise CareBots, erst durch die Berührung mit dem Menschen möglich.

Der Austausch haptischer Signale lässt sich in aktive und passive Berührungen unterteilen. Bei den aktiven Berührungen ist der Agent aktiver Teil in der Berührungsausführung, während er sich bei der passiven Berührung lediglich von dem Menschen berühren lässt. Dabei bringen Berührungen mehrere konzeptuelle Herausforderungen mit sich. Dieses Unterkapitel geht näher auf die Grundlagen der Agenten-Mensch Berührung ein, indem zuerst die konzeptuellen Herausforderungen der Berührungsfähigkeit dargelegt werden und anschließend mittels der aktiven und der passiven Berührung Lösungswege vorgestellt werden.

4.5.1 Herausforderungen

Berührungen zwischen Agenten und Menschen können in mehrerer Hinsicht problematisch sein. Zum einen bedeutet die Berührung die gewollte Kollision mit dem Menschen und zum anderen treten auch hier die Effekte des Unscanny Valeys auf [53]. Das Unscanny Valey tritt auf, da es dem Roboter an den Eigenschaften Weichheit und Wärme mangelt.

Weiterhin bedingt die Berührung die Aufhebung der Intimdistanz zwischen Menschen und Agenten. Die Intimdistanz beschreibt den geduldeten Mindestabstand zum Interaktionspartner und variiert je nach Situation, Beziehungsstatus und Kulturkreis. Diese zu achten ist wichtig für die Erhaltung des Wohlbefindens des Menschen während der Interaktion und bildet damit ein Teil der sozialen Intelligenz des Agenten [88].

4.5.2 Aktive Berührungen

Bei der aktiven Berührung müssen die vorgestellten Problematiken besonders bedacht werden, da das Schadensrisiko durch die Eigenbewegung des Roboters hoch ist. Aus technischer Sicht steht somit besonders die Schadensvermeidung bei der Kollision im Vordergrund. Um diese zu erreichen, ist eine Lösungsmöglichkeit die umfassende Wahrnehmung der unmittelbaren Umgebung. Die Entwicklung von künstlicher Haut [73] ist hierbei ein relevantes Thema. Problematisch ist an dieser Haut jedoch, dass die Verarbeitung der vielen Sensordaten einer sehr hohen Rechenleistung bedarf. Alternativ kann der soziale Agent auch mit begrenzter Kraft oder einem ausreichend weichen und masselosen Körper ausgestattet werden. Ansätze hierfür bietet das Feld der Soft Robotics [82]. Dieses forscht daran, Weichheit und Flexibilität in die Eigenschaften und Fähigkeiten von Robotern zu integrieren. Des Weiteren bietet das Themengebiet der Soft Robotics auch für das Überschreiten des Uncanny Valeys Lösungen. Neben der Weichheit besitzt die Soft Robotik Forschungszweige, welche explizit die Haptik von Lebewesen auf Roboter zu übertragen suchen [82]. Offen bleibt die Frage nach dem Umgang mit der Intimdistanz. Zu dieser benötigt der Agent eine hohe soziale Intelligenz. Diese zu konzeptionieren fällt in das Themengebiet der Psychologie und wird daher in dieser Arbeit nicht weiter behandelt.

4.5.3 Passive Berührungen

Die passive Berührung erlaubt es, die zuvor aufgeführten Problematiken zu umgehen oder sehr stark einzugrenzen. Gleichzeitig enthält sie ein weitaus geringeres Schadenspotenzial. Ein Ansatz ist es, Stellen an dem sozialen Agenten vorzugeben, an denen der Mensch diesen nach Belieben berühren darf. Diese Stellen können beispielsweise optisch markiert sein. Dieser Ansatz bietet gleich mehrere Vorteile. Der wichtigste Vorteil ist, dass die Kollision auf einen einzelnen Bereich verringert werden kann und das Thema der Schadensvermeidung somit wesentlich überschaubarer wird. Es ist hier die Aufgabe

des Entwicklers, Stellen am sozialen Agenten zu wählen, an welchen ein Schadensszenario ausgeschlossen werden kann. Mit der Einschränkung des Berührungsbereiches sinkt ebenfalls die Anzahl an Sensoren und damit die Rechenleistung für die Wahrnehmung der Berührung sehr stark. Selbiger Vorteil gilt für das Überschreiten des Unscanny Valleys. Auch hier muss die Haptik nur in dem vorgegebenen Bereich der Berührung an ein Lebewesen angepasst werden. Weiterhin fällt die Problematik der Wahrung der Intimdistanz weg, da es nun der Mensch selbst ist, der diese überbrückt. Diese Vereinfachungen gegenüber der aktiven Berührung sorgen dafür, dass Soziale Agenten mit der Funktion der passiven Berührung bereits heutzutage in der Praxis etabliert sind [92]. In Bezug auf die Umsetzung eignen sich für die Wahrnehmung der Berührung Drucksensoren oder auch kapazitive Sensoren, welche die unmittelbare Nähe eines Menschen detektieren [72]. Ist der Körper des sozialen Agenten lediglich auf einem Bildschirm dargestellt, könnte vereinfacht auch die Touchfunktion des Bildschirmes zur passiven Berührung genügen.

5 Architekturen sozialer Agenten

Wie im vorhergehenden Kapitel dargelegt, bestehen die Kernfunktionen eines sozialen Agenten aus dem sozialen Wahrnehmen und Handeln innerhalb seiner Umgebung. Diese Kernfunktionen implizieren einen dritten Funktionsbereich des sozialen Agenten: die Verarbeitung von Daten. Damit aus der Wahrnehmung eine Handlung abgeleitet werden kann, benötigt der soziale Agent innere Datenanalysen und Steuerungsprozesse. Die wahrgenommenen Informationen müssen ausgewertet und aufgrund ihrer Aussagen müssen Handlungsentscheidungen getroffen werden. Für diese Prozesse lassen sich verschiedene Architekturüberlegungen anstellen. Diese werden in diesem Kapitel näher beleuchtet. Für die Betrachtung der Architektur sozialer Agenten wird zuerst auf den allgemeinen Aufbau dieser eingegangen. Anschließend werden Prozessmodelle für die Strukturierung konkreter Funktionalitäten vorgestellt. Abschließend werden die gängigsten Bestandteile eines sozialen Agenten anhand einer Beispielarchitektur veranschaulicht.

5.1 Allgemeiner Aufbau

Es gibt viele verschiedene Ansätze, die allgemeine Architektur des Agenten umzusetzen. Einige Ansätze treten sehr häufig auf und werden daher in diesem Unterkapitel vorgestellt. Zum einen ist dies die Unterscheidung in statefull und stateless Architekturen und zum anderen ist dies der Aufbau sozialer Agenten als verteiltes System.

Statefull und stateless Architekturen

Agenten können für eine erste Unterteilung in reaktive und adaptive Agenten unterschieden werden. Reaktive Agenten verfügen über eine stateless Architektur, während adaptive über eine statefull Architektur implementiert werden. Eine Statefull Architecture zeichnet aus, dass die Anwendung Informationen aus vorherigen Ausführungen speichert und sich somit über die Zeit verändert. Eine Stateless Architecture hingegen ist dadurch charakterisiert, dass sie keine neuen Informationen speichert und damit immer wieder von der selben Ausgangsposition startet. Die Umsetzung von stateless Agenten gestaltet sich meist einfacher, wobei die Anwendungsmöglichkeiten für sich genommen recht eingeschränkt sind. Dies ändert sich, wenn die reaktive Anwendung Teil eines verteilten Systems ist. Erhaltene Informationen können dann an eine statefull Anwendung weitergeleitet und dort abgespeichert oder wieder aufgerufen werden.

Der Agent als verteiltes System

Die Funktionen, die ein sozialer Agent ausführt, können sehr schnell sehr aufwändig werden. Dabei gestaltet es sich jedoch schwierig, hohe Rechenleistung kostengünstig in einen kleinen, für den Agenten angemessenen Formfaktor zu bringen. Daher zeichnet sich ein deutlicher Trend ab, diese Funktionen „As a Service“ von anderen Rechnern zu beziehen. Der Agent leitet die zu verarbeitenden Informationen an diese Rechner weiter und erhält von ihnen das gewünschte Ergebnis. Genauso verhält es sich mit dem Speichern und Abrufen von Daten. Die Informationen werden extern in großen Datenbanken verwaltet und auf Anfrage des Agenten zur Verfügung gestellt. Dieser Ansatz kann für sämtliche Aspekte des Agenten angewandt werden. Der Agent für sich genommen schrumpft damit zum Front-End der Anwendungen zusammen.

Der Nachteil eines solchen Outsourcen der Ressourcen der Roboter ist jedoch die Abhängigkeit der Roboter von dem Service Provider und der Internetverbindung. Stehen die benötigten Ressourcen einmal nicht zur Verfügung, kann der Roboter seine Aufgaben nicht mehr ausführen. Je nach Aufgabenfeld und Umfang der Auslagerung, kann der soziale Agent bei mangelhafter Internetübertragung oder der Produktaufgabe von Seiten des Providers seinen Nutzen verlieren.

5.2 Prozessmodelle

Der soziale Agent kann die Datenanalyse- und Steuerungsprozesse auf verschiedenste Weisen ausführen. Um eine Orientierung zu geben, wie diese Prozesse ausgeführt werden können, werden im Folgenden drei Modelle vorgestellt. Das erste Modell bietet eine allgemeine Organisationsform sämtlicher Prozesse nach zeitlichen Kriterien. Das zweite Modell dient der internen Darstellung und Verarbeitung von Emotionen. Das dritte Modell geht auf die Umsetzung des Verhaltens des Agenten ein. Wie in Unterabschnitt 3.4 dargestellt, ist der dynamische Umgang mit Emotionen und Verhalten ein wichtiger Bestandteil der Interaktion mit dem Menschen. Es wird zuerst die Ablauforganisation vorgestellt und anschließend auf die Emotions- sowie die Verhaltensmodellation eingegangen.

Ablauforganisation

Besitzt der soziale Agent die Möglichkeit Daten abzuspeichern, können sowohl die Datenanalyse als auch die Steuerungsprozesse organisatorisch in mehrere Zeitbereiche un-

terteilt werden [65]. Diese Zeitbereiche teilen sich auf in die kurzfristigen, mittelfristigen und langfristigen Betrachtungsräume der Wahrnehmung und Handlung. Der kurzfristige Bereich umfasst die Zeitspanne der nächsten Sekunden. In diesen fällt die Wahrnehmung und Auswertung aktueller Sensordaten und die richtige und synchrone Ausführung der aktuellen Handlung. Der mittelfristige Bereich betrachtet die Zeitspanne der nächsten Stunden oder auch des gesamten Tages. Wahrnehmungen und Handlungen werden hier in den Kontext der letzten Stunden gesetzt. Es wird damit möglich, eine konsistentere Wahrnehmung von dem Menschen zu erhalten und Handlungen zu organisieren. Der langfristige Bereich ist von den Zielen des sozialen Agenten geprägt. In diesen Bereich fallen Charakteranalysen des Interaktionspartners, Beziehungsmodelle und die allgemeine Planung zum Erreichen des Ziels des Agenten. Je nach Anwendungsfall können die letzten beiden Stufen auch zu einer zusammengefasst werden.

Emotionsmodellation

Bei den Emotionen des Agenten ist ausschlaggebend, ob diese als diskrete oder kontinuierliche Werte angegeben werden [25]. Kontinuierliche Emotionswerte bieten potenziell realistischere Möglichkeiten der Emotionsverarbeitung, während sich diskrete Werte einfacher verarbeiten lassen. Weiterhin ist es wichtig für die Glaubwürdigkeit des Agenten, dass dieser immer auch eine neutrale Emotion ausstrahlen kann. Im Folgenden werden zuerst Möglichkeiten zur Darstellung von kontinuierlichen und anschließend von diskreten Emotionswerten vorgestellt.

Kontinuierlichen Emotionswerte lassen sich in mehrdimensionalen Räumen darstellen. Eine unstrukturierte Anordnung der Darstellungsweise ist als n-Dimensionale Koordinate möglich, wobei n die Anzahl der umsetzbaren Grundemotionen ist. Jede dieser Emotionen wird durch eine Koordinate beschrieben. Strukturieren lässt sich die Emotionsdarstellung, indem die Emotionen in ihre grundlegenden Eigenschaften zerlegt und nur diese Eigenschaften als Dimensionen abgebildet werden. Etabliert ist hier die Darstellung mittels der Eigenschaften „*Pleasure-Arousal-Dominance*“ (PAD) nach Mehrabian und Russell [52]. Es ist jedoch auch möglich, die Emotionen in ihrer Abfolge zu strukturieren und beispielsweise einen Emotionskreis vergleichbar mit Plutchiks aufzubauen. Hierzu kann ein Normalkreis in einem zweidimensionalen Koordinatensystem beschrieben werden [25].

Die Verwendung diskreter Emotionswerte ermöglicht nur ein eingeschränktes Abbild des

menschlichen Emotionsspektrums. Aus dem Emotionsspektrum des Menschen werden hierbei spezifische Punkte ausgewählt und dem sozialen Agenten zur Verfügung gestellt. Der soziale Agent verfügt anschließend nur über genau die Emotionen, in welche er wahrgenommene oder geplante Emotionen kategorisch einordnen und ausführen kann. Um ungewöhnliche Sprünge in den Emotionen zu vermeiden, bietet sich die Strukturierung der Emotionspunkte an. Eine geeignete Datenstruktur ist dabei der Graph.

Verhaltensmodellation

Auf kurzfristige Sicht ist das Verhalten des Agenten die Ableitung der nächsten Handlung aufgrund der ihm zu Verfügung stehenden Informationen. Auf langfristige Sicht erzeugt das Verhalten den Charakter des Agenten. Die Veränderung des Verhaltens des Agenten erlaubt es diesem, sich besser an Aufgaben, Situationen und Menschen anzupassen. Um dieses zu erreichen, müssen die Verhaltensmerkmale des Agenten variabel gestaltet werden. Die Verhaltensmerkmale können sich dabei auf die unterschiedlichsten Bereiche erstrecken. Innerhalb der Bereiche können Charakteristiken, wie Ausführungsgeschwindigkeit, Wiederholungsdauer oder auch Intensität, angepasst werden. Welches Verhalten in wie weit zu einer Aufgabe, Situation oder auch einem Menschen passt, kann dabei nur durch Nutzerstudien herausgefunden werden. Eine Übersicht über bisherige Ansätze kann in dem Paper „*A Systematic Review of Adaptivity in Human-Robot Interaction*“ [2] von Ahmad, Mubin und Orlando gefunden werden. Im Folgenden wird darauf eingegangen, wie eine Struktur zur Anpassung der Verhaltensweisen beziehungsweise des Charakters des sozialen Agenten aufgebaut und wie der Charakter des Agenten auf den Nutzer angepasst werden kann.

Wie angesprochen, können durch Kombinationen von Verhaltensmerkmalen Persönlichkeiten simuliert werden. Für die Strukturierung dieser Persönlichkeiten kann eines der vorgestellten Persönlichkeitsmodelle (siehe Unterabschnitt 3.4) herangezogen werden. Um diese auf das Verhalten des Agenten zu übertragen, könnte unter Zuhilfenahme von Nutzerstudien für jeden Parameter des gewählten Modells eine Persönlichkeit des Agenten modelliert werden. Damit wird erreicht, dass das Verhaltensspektrum des sozialen Agenten alle Charakteristika des gewählten Persönlichkeitsmodells abdeckt. Durch Interpolation der Werte der Verhaltenseigenschaften wird es nun möglich, jede Persönlichkeit innerhalb dieses Modells zu realisieren. Der soziale Agent kann das Persönlichkeitsmodell damit vollständig abbilden.

Für die Umsetzung der Verhaltensadaption an Einzelpersonen ist eine einfache Möglichkeit, das Individuum bei Erstkontakt über dessen Persönlichkeit und Werte zu befragen [9]. Die Antworten können anschließend ausgewertet und das Verhalten des Agenten entsprechend justiert werden. Möglich sind jedoch auch dynamische Optimierungsverfahren, nach welchen der soziale Agent seine Persönlichkeitsparameter einstellt. Diese Optimierungen werden auch als ein Lernen des Agenten bezeichnet. Hierzu können Reinforcement Learnings Verfahren, wie beispielsweise das Policy Gradient Reinforcement Learning (PGRL) oder das Q-Learning, eingesetzt werden [80]. Um zu Lernen braucht es Feedback zum ausgeführten Verhalten. Eine Möglichkeit hierzu ist die Emotionserkennung. Mittels dieser kann von der Reaktion des Menschen zeitnah auf die Wirkung der Handlung zurückgeschlossen werden. Es taugen jedoch auch andere Faktoren, wie die Geschwindigkeit für das Erreichen eines Zieles, die Interaktionshäufigkeit oder die Interaktionsdauer, für den Erhalt von Feedback.

5.3 Beispielarchitektur

Abschließend sei hier eine umfängliche Beispielarchitektur eines sozialen Agenten vorgestellt. Bei diesem Beispiel handelt es sich um einen adaptiven sozialen Agenten mit nicht-verteilter Architektur. Inspirationen für diesen Entwurf wurden aus den Veröffentlichungen „*The knowledge level*“ [55] von Newell, „*Operating in a hierarchy of time scales for an always-on relational agent*“ [65] von Rich, Sidner, Nooraei und Coon und „*Tentative Formalization of Human-Agent Interaction for Model-Based Interaction Design*“ [69] von Sakamoto und Tekeuchi genommen. Dieses Beispiel erhebt dabei weder Anspruch auf Vollständigkeit noch auf universelle Gültigkeit. Die zentralen Komponenten der Architektur werden im Folgenden im Einzelnen vorgestellt.

- **Datenanalyse**

In der Datenanalyse werden die aufgenommenen Daten der Wahrnehmungsprozesse zusammengeführt und ausgewertet. Die Datensätze werden mit ihren Redundanzen abgeglichen und nach ihrer Aussagekraft bewertet. Weiterhin werden Datensätze zusammengeführt und ausgewertet, um so weitere Informationen aus ihnen zu gewinnen.

- **Zielvorgaben**

Ausgangspunkt für jede Planung ist ein Ziel. Das Ziel kann dabei aus verschiedenen Punkten bestehen, welche je nach Situation ihre Priorität ändern können.

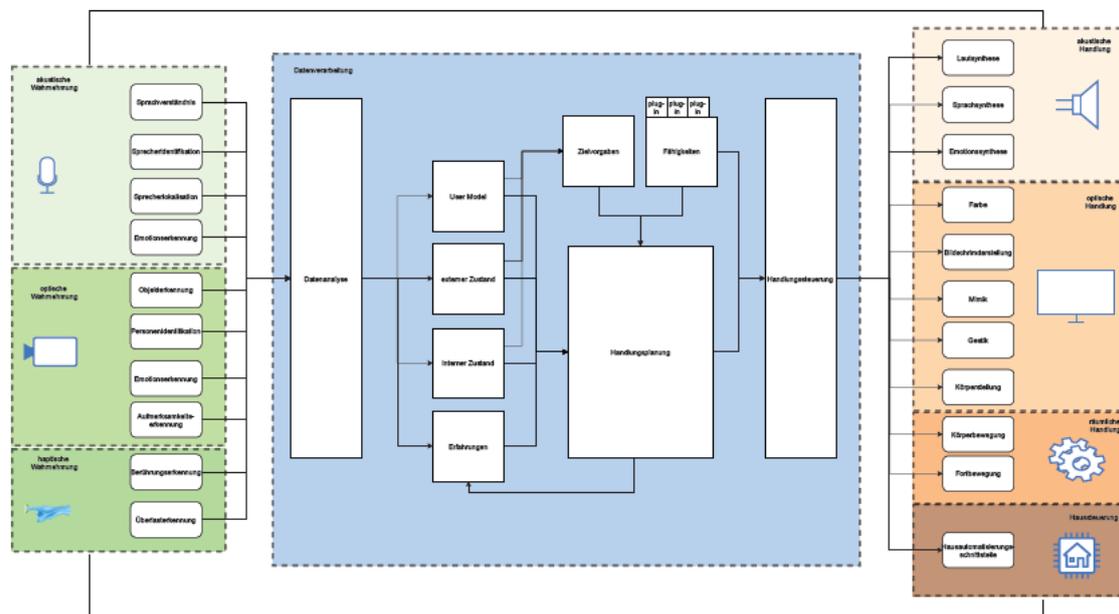


Abbildung 8: Architekturbeispiel eines adaptiven sozialen Agenten

Neben den aufgabenspezifischen Zielen ist es meist sinnvoll allgemeine Ziele vorzugeben, welche beim Erreichen jedes aufgabenspezifischen Ziels berücksichtigt werden sollen. Ein Beispiel für solche allgemeinen Ziele sind Asimov's Grundregeln für funktionssichere Roboter [6].

- **User Model**

Die Kernfähigkeit des sozialen Agenten ist die Interaktion mit dem Menschen. Folglich ist es für den Agenten wichtig, Informationen über diesen zu erhalten und zu speichern. So ist es möglich ein virtuelles Modell des Interaktionspartners, sprich ein User Model, zu schaffen. Je nach Anwendungsgebiet und Informationslage, kann das User Model in seiner Komplexität und Aussagekraft sehr variieren.

- **externer Zustand**

Bei der sozialen Interaktion ist es oft nicht ausreichend, nur die Informationen über den Nutzer in der Handlung zu berücksichtigen. Der Nutzer muss in den Kontext der Umgebung gesetzt werden. Um die Masse an wahrgenommenen Daten über die Umwelt effizient verarbeiten zu können, bietet es sich an, einen aktuellen externen Zustand aus diesen abzuleiten. Dabei werden die umfänglichen Informationen mit den Kriterien vordefinierter Zustände abgeglichen und unter Betrachtung des vorherigen Zustandes der passendste aktuelle Zustand gewählt. Wie aussagekräftig

dieses Modell der Umwelt ist, ist auch hier sehr von der Güte der Wahrnehmungsfähigkeit des Agenten abhängig und daher generell schwierig zu beantworten.

- **interner Zustand**

Neben dem Wissen um die Umwelt des Agenten, muss dieser auch Informationen über sich selbst besitzen, um sinnvolle Planungen aufstellen zu können. Dazu muss der Agent neben einem Modell für den externen auch ein Modell für den internen Zustand besitzen. Weiterhin erlaubt die Berücksichtigung des internen Zustandes die Umsetzung eines komplexeren Charakters, bei dem auch bei reaktiven Handlungen der interne Zustand des Agenten miteinbezogen wird.

- **Erfahrungen**

Des Weiteren verfügt der soziale Agent in diesem Beispiel über die Möglichkeit Erfahrungen zu sammeln. Dies geschieht, indem der soziale Agent seine Planungen und Handlungen mit deren Wirkungen abgleicht. Die Wirkung kann, wie in Abschnitt 5.2 angesprochen, aus verschiedensten Daten abgeleitet werden. Aus der Differenz zwischen Zielvorstellung und Wirkung kann anschließend die Effektivität der Handlung oder der Planung rückgeführt werden. Diese Werte für die Effektivität einer Handlung oder Planungsweise werden abgespeichert und für zukünftige Planungen und Handlungen bereitgestellt.

- **Fähigkeiten**

Neben den aktuellen Informationen über die operative Umwelt des sozialen Agenten benötigt dieser auch Informationen über seine abstrakteren Fähigkeiten. Nicht nur welche Fähigkeiten, sondern auch wie diese Fähigkeiten ausgeführt werden, ist eine kritische Information für die Handlung des Agenten. Weiterhin wird die Idee von Rich, Sidner, Nooraei und Coon aufgegriffen, die Fähigkeiten mittels Software Plugins dynamisch erweitern zu können [65].

- **Handlungsplanung**

In der Handlungsplanung laufen sämtliche Informationen zusammen. Bei der Konzeptionierung der Handlungsplanung ist zu beachten, dass nicht jedes Ereignis im operativen Umfeld des sozialen Agenten planbar sein wird. Weiterhin bauen die Planungen auf virtuellen Modellen auf, deren Aussagekraft immer ungewiss sein wird. Diesbezüglich wird vom sogenannten Ramification Problem gesprochen. Daraus ist abzuleiten, dass mit Fehlern bei der Planung zu rechnen ist. Besonders in kritischen Situationen sollte der soziale Agent somit auch reaktive Handlungen,

spricht Handlungen ohne, oder mit lediglich sehr geringer vorhergehender Planung, ausführen können.

- **Handlungssteuerung**

Die Handlungssteuerung führt die Anweisungen der Handlungsplanung anhand der Vorlagen der zur Verfügung stehenden Fähigkeiten aus. Dazu werden einzelnen Handlungsstränge erstellt, koordiniert und an die spezifischen Handlungsebenen ausgegeben. Bei geringer Komplexität der Komponenten kann die Handlungssteuerung auch mit der Handlungsplanung zusammengeführt werden.

6 Designempfehlungen

In diesem Kapitel wird geschlussfolgert, was aus den vorhergegangenen Betrachtungen in Bezug auf die Konzeptentwürfe sozialer Agenten für das Living Place zu beachten ist. Es wird dabei zuerst auf die generellen Aspekte der Konzeptionierung sozialer Agenten eingegangen. Darauffolgend wird der Bereich der Funktionskonzepte sozialer Agenten beleuchtet.

Generelle Aspekte der Konzeptionierung

Dieses Unterkapitel betrachtet drei generelle Aspekte der Konzeptionierung sozialer Agenten. Zuerst wird auf die natürlichen Stärken und Schwächen sozialer Agenten eingegangen. Anschließend folgen Empfehlungen zur physischen Umsetzung sozialer Agenten. Abschließend wird die Wichtigkeit des nutzenorientierten Entwicklungsvorgehen betont und eine Möglichkeit zur Umsetzung dieser vorgestellt.

Es ist effizient, auf die Stärken sozialer Agenten zu bauen, während die Schwächen konzeptuell in den Hintergrund gestellt werden. Der Antromorphismus erlaubt es, das kreierte Wesen in seiner Erscheinung sehr abstrakt zu halten, ohne dabei die Anerkennung als Interaktionspartner zu verlieren. Diese Stärke kann beispielsweise bei der Gestaltung des Gesichts des Wesens ausgespielt werden. Durch die Existenz eines abstrakten Gesichts ist es möglich, die Kommunikationsfähigkeit des Agenten zu steigern, ohne hierdurch weitere Funktionen zu implizieren. Auch das Kindchenschema sollte in dem Konzept des Agenten berücksichtigt werden. Mit wenig Aufwand kann so die Zuneigung von Seiten des Menschen und dessen Fehlerakzeptanz gegenüber dem Agenten erhöht werden. Weiterhin ist eine Vielzahl oberflächlicher Beziehungen einer geringen Anzahl tieferer Beziehungen zu bevorzugen.

Bei der physikalischen Umsetzung gilt, dass der Grundeindruck des sozialen Agenten allgemein ansprechend sein sollte. Am leichtesten ist es, wenn das Konzept des sozialen Agenten ohne Körper aufgebaut werden kann. Wird dennoch ein Körper benötigt, empfiehlt es sich, diesen sehr einfach zu halten oder auf eine bestehende Plattform zurückzugreifen. Ein beweglicher Körper sollte dabei möglichst klein und masselos gewählt werden, um so das Kollisionsrisiko zu minimieren. Auf die aktive Berührung durch den sozialen Agenten sollte, aufgrund der Risiken, auch bei geringer Körpergröße und -masse verzichtet werden.

Bei der Entwicklung sozialer Agenten muss immer beachtet werden, dass nur ein Test mit dem Nutzer Aufschluss über die Güte der Funktionen und die generelle Erscheinung bietet. Die Konzepte des User Centered Design sollten daher im Entwicklungsprozess berücksichtigt werden. Eine Möglichkeit, diese Tests fortlaufend mit der Entwicklung umzusetzen, könnte sein, eine Varianz der Autonomie für die Steuerung des sozialen Agenten vorzusehen. Ist es möglich die Autonomie des Agenten in einer Spanne von einer detaillierten Fernsteuerung jeglicher Funktionen bis hin zur kompletten Autonomie des Agenten zu variieren, kann dies das regelmäßige Testen und damit die Entwicklung des Agenten erheblich vereinfachen.

Funktionsspezifische Aspekte der Konzeptionierung

Im Weiteren wird konkreter auf die generelle Funktionsgestaltung sozialer Agenten eingegangen. Begonnen wird mit den Herausforderungen der Funktionsumsetzung sozialer Agenten und deren Lösungsansätzen. Anschließend folgen Vorschläge zur Strukturierung und Vereinfachung der Umsetzung der Funktionen.

Bei den Funktionen des Agenten sollte beachtet werden, dass die Illusion der sozialen Präsenz schnell zerstört ist, wenn dieser den Menschen in seinen Funktionen enttäuscht. Die saubere Umsetzung der Funktionen ist daher essenziell für den sozialen Agenten. Hilfreiche Prinzipien können hierbei sein, die Qualität der Funktionen der Quantität vorzuziehen und den Funktionsumfang auf ein möglichst simples und spezifisches Aufgabenfeld einzugrenzen. Um die Enttäuschung des Nutzers zu verhindern, sollte zudem das Erscheinungsbild möglichst wenige grundlegende Funktionen implizieren. Eine gute Möglichkeit hierzu ist die Designneuschöpfung. Wichtig hierbei ist, dass die Erscheinung des Agenten eine Einzigartigkeit erhält und nicht an bereits existente reale oder fiktive Wesen erinnert. Durch diese Einzigartigkeit gibt es für den Mensch kein Vergleichswert, womit sich die Enttäuschungswahrscheinlichkeit verringert. Optimalerweise erstreckt sich die Einzigartigkeit nicht nur auf die Erscheinung sondern auch auf die Funktionen des Agenten.

Ist es hingegen das Entwicklungsziel, einen sozialen Agenten mit weitem Funktionsumfang zu erschaffen, ist zu überlegen, ob dieser modular aufgebaut werden kann. Verschiedene Fachdisziplinen können so unabhängig voneinander und über Jahre hinweg die Funktionen des sozialen Agenten erarbeiten und implementieren. Von der Fachdisziplin der sozialen Agenten wird somit nur eine Plattform mit sozialen Fähigkeiten als Front-End

geschaffen, in welche weitere Funktionen aus anderen Fachbereichen eingehängt werden können.

Weiterhin sollte überlegt werden, den sozialen Agenten etappenweise aufzubauen. Die Arbeitslast wird somit zeitlich besser aufgeteilt. Gleichzeitig können über das Erreichen eines Etappenziels bereits umfängliche Funktionstests und Forschungsversuche ermöglicht werden. Dieser Ansatz kann für einzelne Funktionen, wie auch für die gesamte Architektur des sozialen Agenten, angewandt werden. Bei der Etappierung der Einzel-funktionen bietet es sich beispielsweise an, einen ChatBot zu entwickeln und diesen im Weiteren zu einem VoiceBot auszubauen. In Bezug auf die Gesamtarchitektur kann zunächst ein reaktiver Agenten umgesetzt werden, welcher in weiteren Schritten zu einem adaptiven Agenten ausgebaut wird.

Bei der Betrachtung der konkreten Umsetzung der Kernfunktionen eines sozialen Agenten ist zu sagen, dass es hierzu bereits viel im Netz verfügbare Software existiert. Besonders bei komplexeren Funktionen, wie dem Sprachverständnis oder der Sprachgenerierung, sollten die auf dem Markt verfügbaren Optionen berücksichtigt werden. Eine Eigenentwicklung wird hier in den meisten Fällen überflüssig sein. Des Weiteren ist zu überlegen, ob vorhergegangene Arbeiten im Living Place in die Funktionen des sozialen Agenten integriert werden können. Auf diese Weise kann von bereits verfügbarem Wissen und vorhandenen Funktionen profitiert werden; seien es simple Steuerungsfunktionen des Smart Homes oder das Integrieren komplexerer Funktionen, wie beispielsweise das Erkennen und Tracken von Personen über die Kameras des Living Place. Weiterhin eignen sich die Server des CSTI für die Bereitstellung der benötigten Rechenleistung und des benötigten Speichers des Agenten.

7 Fazit und Ausblick

Fazit

In dieser Arbeit wurden zwei Fragen ausgiebig behandelt. Zum einen wurde sich mit der Frage auseinandergesetzt, was das Themengebiet der sozialen Agenten umfasst, und zum anderen wurde die Frage nach Umsetzungsmöglichkeiten sozialer Agenten bearbeitet.

Die erste Frage wurde bearbeitet, indem verschiedene Begrifflichkeiten eingeordnet, Dienstfelder beschrieben und Produkte vorgestellt wurden. Obwohl sich auf das Anwendungsgebiet des Living Place eingegrenzt wurde, konnten nur die populärsten Aspekte dieser Betrachtungen berücksichtigt werden. Es stellte sich heraus, dass das Themenfeld der sozialen Agenten sehr breit und dabei kaum definiert ist. In dieser Arbeit konnten somit nur die Kerngebiete sozialer Agenten abgedeckt werden. Für den mit dieser Arbeit beabsichtigten Einstieg in das Themengebiet der sozialen Agenten sollten die hier angestellten Betrachtungen jedoch genügen.

Inhaltlich geht aus diesen Betrachtungen hervor, dass sich das Themenfeld der sozialen Agenten stark im Wachstum befindet, wobei sich bisher wenige Konzeptansätze und Definitionen etabliert haben. Das Anwendungspotenzial sozialer Agenten ist, aufgrund der vielfältigen Einsatzbereiche, enorm. Ebenso sind es die Erwartungen an den breiteren Erfolg sozialer Agenten. Die Realität sieht jedoch noch wenig euphorisch aus. Die Forschung an sozialen Agenten ist sehr aktiv, während sich der Erfolg in der Praxis nur auf wenige Anwendungsbereiche erstreckt. Durchsetzen können sich vor allem Plattformen für den Aufbau sozialer Agenten und körperlose Assistenzsysteme. Ansonsten mangelt es den Produkten meist noch an Ausgereiftheit und Wirtschaftlichkeit, um in der Praxis erfolgreich sein zu können.

Die zweite Frage wurde bearbeitet, indem einleitend Überlegungen zu dem allgemeinen Design sozialer Agenten angestellt wurden. Darauf aufbauend wurden die Kernfunktionen eines sozialen Agenten vorgestellt und Architekturansätze besprochen. Den Abschluss bilden die Designempfehlungen, welche aus diesen Betrachtungen resultieren. Wie auch für die Behandlung der ersten Frage, stellte sich auch hier der Umfang und die Unübersichtlichkeit des Themenfeldes als problematisch heraus. Neben dem Fokus auf die technische Sicht der Umsetzung wurde dieser Problematik mit weiteren Eingrenzungen des Themenfeldes entgegengewirkt. So betrachtet diese Arbeit ausschließlich die Kommunikation

als Interaktionsform mit dem Mensch und erarbeitet weiterhin eine Definition der Kernfunktionen des sozialen Agenten, um dessen Funktionsumfang herunterzubrechen. Auch wird sich bei der Umsetzung nicht auf Anwendungsbezüge festgelegt, um so inhaltlich im Allgemeinen zu bleiben und Spezialfälle zu meiden. Die Behandlung der zweiten Frage kann damit ebenfalls keinen Anspruch auf Vollständigkeit erheben. Eine Orientierung bei der Entwicklung sozialer Agenten innerhalb des Living Place sollte diese Arbeit jedoch geben.

Inhaltlich geht aus diesen Betrachtungen hervor, dass es verschiedenste Möglichkeiten gibt, soziale Agenten zu gestalten. Oft existieren redundante Wege auf unterschiedlichen Kommunikationsebenen für die Umsetzung der Kernfunktionen. Diese Redundanzen sollten bei dem Entwurf sozialer Agenten Beachtung finden, da nur so das volle Kommunikationsspektrum des Menschen abgedeckt werden kann. Eine weitere Erkenntnis dieser Arbeit ist, dass das subjektive Empfinden des Menschen ein Schlüsselfaktor für das erfolgreiche Design eines sozialen Agenten darstellt. Dies sollte besonders bei der Konzeptionierung der Fähigkeiten des sozialen Agenten berücksichtigt werden. Funktionen, die über die individuellen Kernfunktionen des sozialen Agenten hinausgehen, sollten vorerst vernachlässigt werden. So wird der Funktionsumfang minimiert, womit das Prinzip der Qualität vor Quantität in der Funktionsumsetzung ausreichend berücksichtigt werden kann.

Ausblick

Mit dieser Arbeit ist ein Grundstein gelegt, um das Themengebiet der sozialen Agenten an der HAW-Hamburg im Living Place zu erforschen. Der nächste Schritt ist das Ersinnen und Auswählen von Konzeptideen speziell für das Living Place. Wichtig hierbei ist es, möglichst einfach umzusetzende Nutzertests während der Entwicklung einzuplanen. Denkbar ist auch ein langfristigeres Projekt für die Entwicklung eines komplexeren sozialen Agenten. Für die Realisierbarkeit ist dabei jedoch zu beachten, dass ein solches Projekt in möglichst unabhängige und handliche Aufgabenpakete unterteilbar sein muss. Nur so kann sichergestellt werden, dass bei dem semesterweisen Wechsel der Studierenden das Projekt abgeschlossen werden kann.

Diese Arbeit wurde am Beispiel des Living Place der HAW-Hamburg geschrieben. Der Großteil der hier abgedeckten Themen ist jedoch auch für den Einsatz sozialer Agenten in anderen Wohneinrichtungen gültig. Weiterhin können viele Aspekte dieser Arbeit

ebenfalls auf soziale Agenten außerhalb des Anwendungsgebiets des Privatwohnraums übertragen werden.

Da diese Arbeit stark auf konkrete Themenbereiche eingegrenzt wurde, blieben zum Abschluss viele Aspekte sozialer Agenten unbehandelt. In wie weit kann mit einem Menschen auf außerhalb der ihm geläufigen Wege kommuniziert werden? Wie kann eine Interaktion mit dem Menschen außerhalb der Kommunikation gestaltet werden? Welche Konzepte sozialer Agenten kommen konkret für das Living Place in Frage? Dies sind Fragen für folgende Arbeiten.

Es bleibt abzuwarten, wie schnell der Siegeszug der sozialen Agenten sein wird. Ob es bei den sozialen Fähigkeiten einer Maschine Grenze gibt und in wie weit eine Maschine die menschliche Interaktion ergänzen oder gar ersetzen kann, muss erst durch zukünftige Forschungen in Erfahrung gebracht werden. Wie auch immer die Antworten auf diese Fragen lauten mögen, es ist nur eine Frage der Zeit bis soziale Agenten fest zu unserem Alltag und unserer Gesellschaft gehören.

Abbildungsverzeichnis

1	Grundriss des Laborbereiches des Living Places (Quelle: https://livingplace.haw-hamburg.de/)	5
2	NAO (Quelle: https://robots.ieee.org/robots/nao/)	11
3	Pepper (Quelle: https://bit.ly/3gwLRcJ)	12
4	Furhat (Quelle: https://www.robotlab.com/store/furhat-robot)	12
5	Jibo (Quelle: https://bit.ly/32yFDAS)	14
6	Vector (Quelle: https://www.sbsupply.de/anki-vector)	14
7	Gatebox (Quelle: https://bit.ly/3egX7He)	15
8	Architekturbeispiel eines adaptiven sozialen Agenten	43

Referenzen

- [1] ACKERMAN, Evan: *Spectrum IEEE Can Amazon Build a Home Robot That Is Useful and Affordable?* 2018. – URL <https://spectrum.ieee.org/automan/robotics/home-robots/amazon-home-robots>. – Zugriffsdatum: 2021-04-16
- [2] AHMAD, Muneeb ; MUBIN, Omar ; ORLANDO, Joanne: A Systematic Review of Adaptivity in Human-Robot Interaction. In: *Multimodal Technologies and Interaction* 1 (2017), Nr. 3. – URL <https://www.mdpi.com/2414-4088/1/3/14>. – ISSN 2414-4088
- [3] ALBERT, Mehrabian: *Nonverbal communication*. 1972
- [4] ALGHAMDI, J. ; ALHARTHI, R. ; ALGHAMDI, R. ; ALSUBAIE, W. ; ALSUBAIE, R. ; ALQAHTANI, D. ; RAMADAN, R. A. ; ALQARNI, L. ; ALSHAMMARI, R.: A Survey On Face Recognition Algorithms. In: *2020 3rd International Conference on Computer Applications Information Security (ICCAIS)*, 2020, S. 1–5

- [5] ANDRE, Elisabeth: *Lässt sich Empathie simulieren? Ansätze zur Erkennung und Generierung empathischer Reaktionen anhand von Computermodellen*. S. 81–105, 01 2014
- [6] ASIMOV, Isaac: Three laws of robotics. In: *Asimov, I. Runaround* (1941)
- [7] BATISH, Rachel: *Voicebot and Chatbot Design: Flexible Conversational Interfaces with Amazon Alexa, Google Home, and Facebook Messenger*. Packt Publishing Ltd, 2018
- [8] BECK, Aryel ; STEVENS, Brett ; BARD, Kim A. ; CAÑAMERO, Lola: Emotional Body Language Displayed by Artificial Agents. In: *ACM Trans. Interact. Intell. Syst.* 2 (2012), März, Nr. 1. – URL <https://doi.org/10.1145/2133366.2133368>. – ISSN 2160-6455
- [9] BENDEL, Oliver: *Handbuch Maschinenethik*. Springer, 2019
- [10] BENNETT, Casey C. ; ŠABANOVIĆ, Selma: Perceptions of Affective Expression in a Minimalist Robotic Face. In: *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction*, IEEE Press, 2013 (HRI '13), S. 81–82. – ISBN 9781467330558
- [11] BETHEL, C. L. ; MURPHY, R. R.: Survey of Non-facial/Non-verbal Affective Expressions for Appearance-Constrained Robots. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 38 (2008), Nr. 1, S. 83–92
- [12] BEVAN, Nigel ; CARTER, James ; HARKER, Susan: ISO 9241-11 Revised: What Have We Learnt About Usability Since 1998? In: KUROSU, Masaaki (Hrsg.): *Human-Computer Interaction: Design and Evaluation*. Cham : Springer International Publishing, 2015, S. 143–151. – ISBN 978-3-319-20901-2
- [13] BICKMORE, Timothy ; CASSELL, Justine: Relational Agents: A Model and Implementation of Building User Trust. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA : Association for Computing Machinery, 2001 (CHI '01), S. 396–403. – URL <https://doi.org/10.1145/365024.365304>. – ISBN 1581133278
- [14] BICKMORE, Timothy W. ; PICARD, Rosalind W.: Establishing and Maintaining Long-Term Human-Computer Relationships. In: *ACM Trans. Comput.-Hum. Interact.* 12 (2005), Juni, Nr. 2, S. 293–327. – URL <https://doi.org/10.1145/1067860.1067867>. – ISSN 1073-0516

- [15] BIUNDO, Susanne ; HÖLLER, Daniel ; SCHATTENBERG, Bernd ; BERCHER, Pascal: Companion-technology: an overview. In: *KI-Künstliche Intelligenz* 30 (2016), Nr. 1, S. 11–20
- [16] BREAZEAL, C. ; KIDD, C.D. ; THOMAZ, A.L. ; HOFFMAN, G. ; BERLIN, M.: Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In: *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2005, S. 708–713
- [17] BREAZEAL, Cynthia: *Designing Sociable Machines*. S. 149–156. In: DAUTENHAHN, Kerstin (Hrsg.) ; BOND, Alan (Hrsg.) ; CAÑAMERO, Lola (Hrsg.) ; EDMONDS, Bruce (Hrsg.): *Socially Intelligent Agents: Creating Relationships with Computers and Robots*. Boston, MA : Springer US, 2002. – URL https://doi.org/10.1007/0-306-47373-9_18. – ISBN 978-0-306-47373-9
- [18] BURKHARDT, Felix ; CAMPBELL, Nick: Emotional speech synthesis. In: *The oxford handbook of affective computing* (2014), S. 286–295
- [19] CAÑAMERO, Lola ; LEWIS, Matthew: Making new “New AI” friends: designing a social robot for diabetic children from an embodied AI perspective. In: *International Journal of Social Robotics* 8 (2016), Nr. 4, S. 523–537
- [20] COLGATE, J E. ; EDWARD, J ; PESHKIN, Michael A. ; WANNASUPHOPRASIT, Witaya: Cobots: Robots for collaboration with human operators. (1996)
- [21] COLLOBERT, Ronan ; WESTON, Jason ; BOTTOU, Léon ; KARLEN, Michael ; KAVUKCUOGLU, Koray ; KUKSA, Pavel: Natural language processing (almost) from scratch. In: *Journal of machine learning research* 12 (2011), Nr. ARTICLE, S. 2493–2537
- [22] CORNEANU, C. A. ; SIMÓN, M. O. ; COHN, J. F. ; GUERRERO, S. E.: Survey on RGB, 3D, Thermal, and Multimodal Approaches for Facial Expression Recognition: History, Trends, and Affect-Related Applications. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38 (2016), Nr. 8, S. 1548–1568
- [23] DECKER, Michael: *Ein Abbild des Menschen: Humanoide Roboter*. S. 41–62. In: BÖLKER, Michael (Hrsg.) ; GUTMANN, Mathias (Hrsg.) ; HESSE, Wolfgang (Hrsg.): *Information und Menschenbild*. Berlin, Heidelberg : Springer Berlin Heidelberg, 2010. – URL https://doi.org/10.1007/978-3-642-04742-8_3. – ISBN 978-3-642-04742-8

- [24] DHAMSANIA, C. J. ; RATANPARA, T. V.: A survey on Human action recognition from videos. In: *2016 Online International Conference on Green Engineering and Technologies (IC-GET)*, 2016, S. 1–5
- [25] DIMITRIEVSKA, Vesna ; ACKOVSKA, Nevena: Behavior Models of Emotion-Featured Robots: A Survey. In: *Journal of Intelligent & Robotic Systems* 100 (2020), Nr. 3, S. 1031–1053
- [26] EBERT, Helmut ; BECKER, Joachim H.: *Formen der Kommunikation*. S. 33–42. In: *Praxishandbuch berufliche Schlüsselkompetenzen: 50 Handlungskompetenzen für Ausbildung, Studium und Beruf*. Berlin, Heidelberg : Springer Berlin Heidelberg, 2018. – URL https://doi.org/10.1007/978-3-662-54925-4_5. – ISBN 978-3-662-54925-4
- [27] EKMAN, Paul: Basic emotions. In: *Handbook of cognition and emotion* 98 (1999), Nr. 45-60, S. 16
- [28] EMBGEN, S. ; LUBER, M. ; BECKER-ASANO, C. ; RAGNI, M. ; EVERS, V. ; ARRAS, K. O.: Robot-specific social cues in emotional body language. In: *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, 2012, S. 1019–1025
- [29] Sicherheit von Maschinen - Allgemeine Gestaltungsleitsätze - Risikobeurteilung und Risikominderung / International Organization for Standardization. Geneva, CH, März 2011. – Standard
- [30] ERP, J. B. F. V. ; TOET, A.: How to Touch Humans: Guidelines for Social Agents and Robots That Can Touch. In: *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 2013, S. 780–785
- [31] EYSSEL, Friederike ; KUCHENBRANDT, Dieta ; BOBINGER, Simon: Effects of Anticipated Human-Robot Interaction and Predictability of Robot Behavior on Perceptions of Anthropomorphism. In: *Proceedings of the 6th International Conference on Human-Robot Interaction*. New York, NY, USA : Association for Computing Machinery, 2011 (HRI '11), S. 61–68. – URL <https://doi.org/10.1145/1957656.1957673>. – ISBN 9781450305617
- [32] FORBES, Chelsea ; KLEIN, Holly ; PANNU, Livleen: How Facial Recognition Systems Work.

- [33] GATEBOX: *Startpage*. 2021. – URL <https://www.gatebox.ai/en/>. – Zugriffsdatum: 2021-04-16
- [34] GREEFF, Joachim de ; NOLFI, Stefano: *Evolution of Implicit and Explicit Communication in Mobile Robots*. S. 179–214. In: NOLFI, Stefano (Hrsg.) ; MIROLLI, Marco (Hrsg.): *Evolution of Communication and Language in Embodied Agents*. Berlin, Heidelberg : Springer Berlin Heidelberg, 2010. – URL https://doi.org/10.1007/978-3-642-01250-1_11. – ISBN 978-3-642-01250-1
- [35] HANSON, David: *Spectrum IEEE Why We Should Build Humanlike Robots*. 2011. – URL <https://spectrum.ieee.org/automaton/robotics/humanoids/why-we-should-build-humanlike-robots>. – Zugriffsdatum: 2021-04-16
- [36] HEGEL, Frank ; MUHL, Claudia ; WREDE, Britta ; HIELSCHER-FASTABEND, Martina ; SAGERER, Gerhard: Understanding Social Robots. In: *2009 Second International Conferences on Advances in Computer-Human Interactions*, 2009, S. 169–174
- [37] HWANG, Tim ; PEARCE, Ian ; NANIS, Max: Socialbots: Voices from the Fronts. In: *Interactions* 19 (2012), März, Nr. 2, S. 38–45. – URL <https://doi.org/10.1145/2090150.2090161>. – ISSN 1072-5520
- [38] IEEE: *Jibp*. 2014. – URL <https://robots.ieee.org/robots/jibo/>. – Zugriffsdatum: 2021-04-16
- [39] IEEE: *Vector*. 2018. – URL <https://robots.ieee.org/robots/vector/>. – Zugriffsdatum: 2021-04-16
- [40] JANOWSKI, Kathrin ; RITSCHER, Hannes ; LUGRIN, Birgit ; ANDRÉ, Elisabeth: *Sozial interagierende Roboter in der Pflege*. S. 63–87. In: BENDEL, Oliver (Hrsg.): *Pflegeroboter*. Wiesbaden : Springer Fachmedien Wiesbaden, 2018. – URL https://doi.org/10.1007/978-3-658-22698-5_4. – ISBN 978-3-658-22698-5
- [41] KĚPUSKA, Veton ; BOHOUTA, Gamal: Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). In: *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, 2018, S. 99–103
- [42] KRÄMER, Nicole C. ; BENTE, Gary: Personalizing e-learning. The social effects of pedagogical agents. In: *Educational Psychology Review* 22 (2010), Nr. 1, S. 71–87

- [43] KULKARNI, P. ; MAHABALESHWARKAR, A. ; KULKARNI, M. ; SIRSIKAR, N. ; GADGIL, K.: Conversational AI: An Overview of Methodologies, Applications Future Scope. In: *2019 5th International Conference On Computing, Communication, Control And Automation (ICCUBEA)*, 2019, S. 1–7
- [44] LACEY, Cherie ; CAUDWELL, Catherine: Cuteness as a 'dark Pattern' in Home Robots. In: *Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction*, IEEE Press, 2019 (HRI '19), S. 374–381. – ISBN 9781538685556
- [45] LEITE, Iolanda ; MARTINHO, Carlos ; PAIVA, Ana: Social robots for long-term interaction: a survey. In: *International Journal of Social Robotics* 5 (2013), Nr. 2, S. 291–308
- [46] LEVIATHAN, Yaniv ; MATIAS, Yossi: Google Duplex: an AI system for accomplishing real-world tasks over the phone. (2018)
- [47] LOPER, Matthew M. ; KOENIG, Nathan P. ; CHERNOVA, Sonia H. ; JONES, Chris V. ; JENKINS, Odest C.: Mobile Human-Robot Teaming with Environmental Tolerance. In: *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction*. New York, NY, USA : Association for Computing Machinery, 2009 (HRI '09), S. 157–164. – URL <https://doi.org/10.1145/1514095.1514124>. – ISBN 9781605584041
- [48] LUCK, Kai von: *Creative Space for Technical Innovations*. 2019. – URL <https://csti.haw-hamburg.de/>. – Zugriffsdatum: 2021-04-16
- [49] LUCK, Kai von: *Living Place*. 2019. – URL <https://livingplace.haw-hamburg.de/>. – Zugriffsdatum: 2021-04-16
- [50] MAHADEV, Nadimpalli ; AUSTIN, Kevin B.: Sound Localization by Robot Using Inter-Aural Time Differences. In: *J. Comput. Sci. Coll.* 30 (2015), April, Nr. 4, S. 50–56. – ISSN 1937-4771
- [51] MAO, R. ; LI, G. ; HILDRE, H. P. ; ZHANG, H.: A Survey of Eye Tracking in Automobile and Aviation Studies: Implications for Eye-Tracking Studies in Marine Operations. In: *IEEE Transactions on Human-Machine Systems* 51 (2021), Nr. 2, S. 87–98
- [52] MEHRABIAN, Albert ; RUSSELL, James A.: *An approach to environmental psychology*. the MIT Press, 1974

- [53] MORI, Masahiro: *Spectrum IEEE The Uncanny Valley: The Original Essay by Masahiro Mori*. 2012. – URL <https://spectrum.ieee.org/automaton/robotics/humanoids/the-uncanny-valley>. – Zugriffsdatum: 2021-04-16
- [54] MUNEA, T. L. ; JEMBRE, Y. Z. ; WELDEGEBRIEL, H. T. ; CHEN, L. ; HUANG, C. ; YANG, C.: The Progress of Human Pose Estimation: A Survey and Taxonomy of Models Applied in 2D Human Pose Estimation. In: *IEEE Access* 8 (2020), S. 133330–133348
- [55] NEWELL, Allen: The knowledge level. In: *Artificial intelligence* 18 (1982), Nr. 1, S. 87–127
- [56] ONYEULO, Eva B. ; GANDHI, Vaibhav: What Makes a Social Robot Good at Interacting with Humans? In: *Information* 11 (2020), Nr. 1. – URL <https://www.mdpi.com/2078-2489/11/1/43>. – ISSN 2078-2489
- [57] PAN, M. K. X. J. ; CHOI, S. ; KENNEDY, J. ; MCINTOSH, K. ; ZAMORA, D. C. ; NIEMEYER, G. ; KIM, J. ; WIELAND, A. ; CHRISTENSEN, D.: Realistic and Interactive Robot Gaze. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, S. 11072–11078
- [58] PAWAR, Rupali V. ; JALNEKAR, Rajesh M. ; CHITODE, Janardan S.: Review of various stages in speaker recognition system, performance measures and recognition toolkits. In: *Analog Integrated Circuits and Signal Processing* 94 (2018), Nr. 2, S. 247–257
- [59] PEPLAU, Hildegard E.: Peplau’s theory of interpersonal relations. In: *Nursing science quarterly* 10 (1997), Nr. 4, S. 162–167
- [60] PLUTCHIK, Robert: The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. In: *American scientist* 89 (2001), Nr. 4, S. 344–350
- [61] POLLMANN, Kathrin ; RUFF, Christopher ; VETTER, Kevin ; ZIMMERMANN, Gottfried: Robot vs. Voice Assistant: Is Playing with Pepper More Fun than Playing with Alexa? In: *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. New York, NY, USA : Association for Computing Machinery, 2020 (HRI ’20), S. 395–397. – URL <https://doi.org/10.1145/3371382.3378251>. – ISBN 9781450370578

- [62] QI, Xianyu ; WANG, Wei ; GUO, Lei ; LI, Mingbo ; ZHANG, Xiaoyu ; WEI, Ran: Building a plutchik's wheel inspired affective model for social robots. In: *Journal of Bionic Engineering* 16 (2019), Nr. 2, S. 209–221
- [63] RAO, K S. ; KOOLAGUDI, Shashidhar G.: *Robust emotion recognition using spectral and prosodic features*. Springer Science & Business Media, 2013
- [64] RATHEE, Kovid: *Meet Google Meena*. 2022. – URL <https://towardsdatascience.com/meet-google-meena-11c26849e12a>. – Zugriffsdatum: 2021-04-16
- [65] RICH, Charles ; SIDNER, Candace L. ; NOORAEI, Bahador ; COON, William: Operating in a hierarchy of time scales for an always-on relational agent. In: *Workshop on Real-Time Conversations with Virtual Agents, Santa Cruz, CA (September 2012)*, 2012
- [66] ROBOTICS, Furhat: *Startpage*. 2021. – URL <https://furhatrobotics.com/>. – Zugriffsdatum: 2021-04-16
- [67] ROCCAS, Sonia ; SAGIV, Lilach ; SCHWARTZ, Shalom H. ; KNAFO, Ariel: The big five personality factors and personal values. In: *Personality and social psychology bulletin* 28 (2002), Nr. 6, S. 789–801
- [68] ROHM, Robert A.: What is DISC. In: *It is a powerful way to understand people and their personality types* (2013)
- [69] SAKAMOTO, Takafumi ; TAKEUCHI, Yugo: *Tentative Formalization of Human-Agent Interaction for Model-Based Interaction Design*. S. 319–320. In: *Proceedings of the 7th International Conference on Human-Agent Interaction*. New York, NY, USA : Association for Computing Machinery, 2019. – URL <https://doi.org/10.1145/3349537.3352806>. – ISBN 9781450369220
- [70] SCHEUTZ, Matthias: Useful Roles of Emotions in Artificial Agents: A Case Study from Artificial Life. In: *Proceedings of AAAI 2004*, AAAI Press, 2004, S. 31–40
- [71] SCHULLER, Björn W.: Speech Emotion Recognition: Two Decades in a Nutshell, Benchmarks, and Ongoing Trends. In: *Commun. ACM* 61 (2018), April, Nr. 5, S. 90–99. – URL <https://doi.org/10.1145/3129340>. – ISSN 0001-0782

- [72] SEFIDGAR, Y. S. ; MACLEAN, K. E. ; YOHANAN, S. ; VAN DER LOOS, H. F. M. ; CROFT, E. A. ; GARLAND, E. J.: Design and Evaluation of a Touch-Centered Calming Interaction with a Social Robot. In: *IEEE Transactions on Affective Computing* 7 (2016), Nr. 2, S. 108–121
- [73] SHEIN, Esther: Softening up Robots. In: *Commun. ACM* 63 (2020), November, Nr. 12, S. 12–14. – URL <https://doi.org/10.1145/3427942>. – ISSN 0001-0782
- [74] SIDNER, Candace ; BICKMORE, Timothy ; RICH, Charles ; BARRY, Barbara ; RING, Lazlo ; BEHROOZ, Morteza ; SHAYGANFAR, Mohammad: An always-on companion for isolated older adults. In: *14th Annual SIGdial meeting on discourse and dialogue*, 2013
- [75] SOFTBANK: *Nao*. 2021. – URL <https://www.softbankrobotics.com/emea/en/nao>. – Zugriffsdatum: 2021-04-16
- [76] SOFTBANK: *Pepper*. 2021. – URL <https://www.softbankrobotics.com/emea/en/pepper>. – Zugriffsdatum: 2021-04-16
- [77] STEPHENS-FRIPP, Benjamin ; NAGHDY, Fazel ; STIRLING, David ; NAGHDY, Gols-hah: Automatic affect perception based on body gait and posture: a survey. In: *International Journal of Social Robotics* 9 (2017), Nr. 5, S. 617–641
- [78] SUN, Z. ; CHEN, J. ; LIANG, C. ; RUAN, W. ; MUKHERJEE, M.: A Survey of Multiple Pedestrian Tracking Based on Tracking-by-Detection Framework. In: *IEEE Transactions on Circuits and Systems for Video Technology* (2020), S. 1–1
- [79] TABET, Y. ; BOUGHAZI, M.: Speech synthesis techniques. A survey. In: *International Workshop on Systems, Signal Processing and their Applications, WOSSPA*, 2011, S. 67–70
- [80] TAPUS, Adriana ; ȚĂPUȘ, Cristian ; MATARIĆ, Maja J.: User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. In: *Intelligent Service Robotics* 1 (2008), Nr. 2, S. 169–183
- [81] THUN, Schulz: F. von (1981): Miteinander reden. In: *Reinbek bei Hamburg*
- [82] UENO, A. ; HLAVÁČ, V. ; MIZUUCHI, I. ; HOFFMANN, M.: Touching a Human or a Robot? Investigating Human-likeness of a Soft Warm Artificial Hand. In: *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 2020, S. 14–20

- [83] VAGAŠ, M. ; GALAJDOVÁ, A. ; ŠIMŠÍK, D.: Techniques for Secure Automated Operation with Cobots Participation. In: *2020 21th International Carpathian Control Conference (ICCC)*, 2020, S. 1–4
- [84] VAN ERP, Jan B. ; KYUNG, Ki-Uk ; KASSNER, Sebastian ; CARTER, Jim ; BREWSTER, Stephen ; WEBER, Gerhard ; ANDREW, Ian: Setting the standards for haptic and tactile interactions: ISO's work. In: *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications* Springer (Veranst.), 2010, S. 353–358
- [85] VANDERBORGHT, Bram: *Spectrum IEEE Robotic Dreams, Robotic Realities: Why Is It So Hard to Build Profitable Robot Companies?* 2019. – URL https://spectrum.ieee.org/automaton/robotics/industrial-robots/robotic-dreams-robotic-realities?utm_source=robots.ieee.org. – Zugriffsdatum: 2021-04-16
- [86] Agentensysteme in der Automatisierungstechnik - Grundlagen / Verein Deutscher Ingenieure. DE, Mai 2018. – Standard
- [87] VOGT, Thurid ; ANDRÉ, Elisabeth: Comparing feature sets for acted and spontaneous speech in view of automatic emotion recognition. In: *2005 IEEE International Conference on Multimedia and Expo IEEE* (Veranst.), 2005, S. 474–477
- [88] WALTERS, M. L. ; DAUTENHAHN, K. ; KOAY, K. L. ; KAOURI, C. ; BOEKHORST, R. ; NEHANIV, C. ; WERRY, I. ; LEE, D.: Close encounters: spatial distances between people and a robot of mechanistic appearance. In: *5th IEEE-RAS International Conference on Humanoid Robots, 2005.*, 2005, S. 450–455
- [89] XIA, Z. ; LEI, Q. ; YANG, Y. ; ZHANG, H. ; HE, Y. ; WANG, W. ; HUANG, M.: Vision-Based Hand Gesture Recognition for Human-Robot Collaboration: A Survey. In: *2019 5th International Conference on Control, Automation and Robotics (ICCAR)*, 2019, S. 198–205
- [90] YASHAS, J. ; SHIVAKUMAR, G.: Hand Gesture Recognition: A Survey. In: *2019 International Conference on Applied Machine Learning (ICAML)*, 2019, S. 3–8
- [91] YU, Dong ; DENG, Li: *AUTOMATIC SPEECH RECOGNITION*. Springer, 2016
- [92] YU, Ruby ; HUI, Elsie ; LEE, Jenny ; POON, Dawn ; NG, Ashley ; SIT, Kitty ; IP, Kenny ; YEUNG, Fannie ; WONG, Martin ; SHIBATA, Takanori u. a.: Use of a therapeutic, socially assistive pet robot (PARO) in improving mood and stimulating

social interaction and communication for people with dementia: Study protocol for a randomized controlled trial. In: *JMIR research protocols* 4 (2015), Nr. 2, S. e4189

Erklärung zur selbstständigen Bearbeitung einer Abschlussarbeit

Gemäß der Allgemeinen Prüfungs- und Studienordnung ist zusammen mit der Abschlussarbeit eine schriftliche Erklärung abzugeben, in der der Studierende bestätigt, dass die Abschlussarbeit „— bei einer Gruppenarbeit die entsprechend gekennzeichneten Teile der Arbeit [(§ 18 Abs. 1 APSO-TI-BM bzw. § 21 Abs. 1 APSO-INGI)] — ohne fremde Hilfe selbständig verfasst und nur die angegebenen Quellen und Hilfsmittel benutzt wurden. Wörtlich oder dem Sinn nach aus anderen Werken entnommene Stellen sind unter Angabe der Quellen kenntlich zu machen.“

Quelle: § 16 Abs. 5 APSO-TI-BM bzw. § 15 Abs. 6 APSO-INGI

Dieses Blatt, mit der folgenden Erklärung, ist nach Fertigstellung der Abschlussarbeit durch den Studierenden auszufüllen und jeweils mit Originalunterschrift als letztes Blatt in das Prüfungsexemplar der Abschlussarbeit einzubinden.

Eine unrichtig abgegebene Erklärung kann -auch nachträglich- zur Ungültigkeit des Studienabschlusses führen.

Erklärung zur selbstständigen Bearbeitung der Arbeit

Hiermit versichere ich,

Name: _____

Vorname: _____

dass ich die vorliegende Seminararbeit – bzw. bei einer Gruppenarbeit die entsprechend gekennzeichneten Teile der Arbeit – mit dem Thema:

Soziale Agenten für den Bereich einer modernen Wohnumgebung

ohne fremde Hilfe selbständig verfasst und nur die angegebenen Quellen und Hilfsmittel benutzt habe. Wörtlich oder dem Sinn nach aus anderen Werken entnommene Stellen sind unter Angabe der Quellen kenntlich gemacht.

- die folgende Aussage ist bei Gruppenarbeiten auszufüllen und entfällt bei Einzelarbeiten -

Referenzen

Die Kennzeichnung der von mir erstellten und verantworteten Teile der Seminararbeit ist erfolgt durch:

<hr/>	<hr/>	
Ort	Datum	Unterschrift im Original