

# Bachelorarbeit

Alexander Polter

Change Me: Entwicklung eines konfliktbasierten  
Companions für das Smart Home



Alexander Polter

# Change Me: Entwicklung eines konfliktbasierten Companions für das Smart Home

Bachelorarbeit eingereicht im Rahmen der Bachelorprüfung  
im Studiengang *Bachelor of Science Angewandte Informatik*  
am Department Informatik  
der Fakultät Technik und Informatik  
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuender Prüfer: Dr. Susanne Draheim  
Zweitgutachter: Prof. Dr. Kai von Luck

Eingereicht am: 13.05.2026



**Alexander Polter**

**Thema der Arbeit**

Change Me: Entwicklung eines konfliktbasierten Companions für das Smart Home

**Stichwörter**

künstlicher Companion, konfliktorientierter Companion, konfliktorientiertes Design, Konflikt, Mensch-Computer-Interaktion, Entwicklung von Companion, large language module, Smart Home

**Kurzzusammenfassung**

Diese Arbeit untersucht das Konzept konfliktorientierter künstlicher Companions und entwickelt einen entsprechenden Prototyp für ein Smart Home. Ausgangspunkt ist die Beobachtung, dass Konflikte – entgegen ihrer häufig negativen Bewertung – auch produktive Effekte in zwischenmenschlichen Beziehungen haben können. Auf Basis theoretischer Grundlagen zu Companion-Systemen und konfliktorientiertem Design wird ein Companion entworfen, der Nutzer aktiv bei der Verfolgung selbst gesetzter Ziele unterstützt, indem er in Konfliktsituationen argumentiert, widerspricht und Kompromisse anbietet. Die prototypische Umsetzung zeigt, dass ein solcher Ansatz technisch realisierbar ist und Konflikte als wertvolles Element zur Förderung von Reflexion und Entscheidungsqualität in der Mensch-Maschine-Interaktion dienen können.

**Alexander Polter**

**Title of Thesis**

Change Me: Development of a Conflict-Based Companion for the Smart Home

**Keywords**

artificial companion, conflict-oriented companion, conflict-oriented design, conflict, human-computer-interaction, companion development, large language module, Smart Home

**Abstract**

---

This thesis explores the concept of conflict-oriented artificial companions and develops a corresponding prototype for a Smart Home. It is based on the observation that conflicts—contrary to their often negative perception—can also have productive effects in interpersonal relationships. Building on theoretical foundations of companion systems and conflict-oriented design, a companion is designed to actively support users in pursuing self-defined goals by arguing, challenging, and offering compromises in conflict situations. The prototypical implementation demonstrates that such an approach is technically feasible and that conflict can serve as a valuable element for fostering reflection and improving decision-making in human-machine interaction.

# Inhaltsverzeichnis

Abbildungsverzeichnis	ix
Abkürzungen	xi
<b>1 Einleitung</b>	<b>1</b>
<b>2 Theoretische Grundlage</b>	<b>3</b>
2.1 Definition für künstliche Companions . . . . .	3
2.1.1 Sozialer Agent . . . . .	4
2.1.2 Adaptive Design . . . . .	6
2.1.3 Engaging Design . . . . .	6
2.1.4 Emotionale Bindung des Nutzer an den Agenten . . . . .	6
2.2 Zentrale Designaspekte für künstliche Companions . . . . .	7
2.2.1 Kommunikationsfähigkeit . . . . .	7
2.2.2 Adaptivity . . . . .	8
2.2.3 Engagement . . . . .	9
2.2.4 Persönlichkeit . . . . .	10
2.2.5 Verkörperung und Erscheinungsbild . . . . .	11
2.3 Konflikt . . . . .	11
2.4 Smart Home . . . . .	12
<b>3 Konfliktorientiertes Design</b>	<b>14</b>
<b>4 Beispielhafte Umsetzung eines konfliktorientierten Companions</b>	<b>15</b>
4.1 Design . . . . .	15
4.1.1 Generelle Idee und Beschreibung des Nutzungsszenarios . . . . .	15
4.1.2 Design-Entscheidungen für den künstlichen Companion . . . . .	16
4.1.3 Ethische Aspekte für das Design des künstlichen Companions . . . . .	18
4.1.4 Systemanforderungen . . . . .	20
4.1.5 Technisches Design . . . . .	22

4.2	Implementierung . . . . .	28
4.2.1	Funktionsweise des Companions . . . . .	28
4.2.2	LLM-Nutzung im Companion . . . . .	33
4.2.3	Aussehen des Companions . . . . .	35
4.2.4	Kommunikation . . . . .	36
4.2.5	Verhalten des Companions . . . . .	36
4.3	Evaluation . . . . .	38
4.3.1	Vorgehensweise der Evaluation . . . . .	38
4.3.2	Auswertung der Evaluation . . . . .	39
<b>5</b>	<b>Zusammenfassung und Ausblick</b>	<b>42</b>
	<b>Literatur</b>	<b>44</b>
	<b>A Anhang</b>	<b>49</b>
A.1	Verwendete Hilfsmittel . . . . .	49
A.2	Prompts . . . . .	49
A.2.1	Prompt: agrees . . . . .	49
A.2.2	Prompt: goodbye . . . . .	50
A.2.3	Prompt: compromise . . . . .	50
A.2.4	Prompt: address . . . . .	51
A.2.5	Prompt: argue . . . . .	52
A.3	Test Konversationen . . . . .	53
A.3.1	Ziel ins Bett gehen: Der Nutzer gibt nach . . . . .	53
A.3.2	Ziel ins Bett gehen: Der Companion gibt nach . . . . .	54
A.3.3	Ziel ins Bett gehen: Es wird ein Kompromiss erzielt . . . . .	54
A.3.4	Ziel die Wohnugn verlassen: Der Nutzer gibt nach . . . . .	55
A.3.5	Ziel die Wohnugn verlassen: Der Companion gibt nach . . . . .	55
A.3.6	Ziel die Wohnugn verlassen: Es wird ein Kompromiss erzielt . . . . .	56
A.3.7	Ziel die Fenster öffnen: Der Nutzer gibt nach . . . . .	56
A.3.8	Ziel die Fenster öffnen: Der Companion gibt nach . . . . .	56
A.3.9	Ziel die Fenster öffnen: Es wird ein Kompromiss erzielt . . . . .	57
	<b>Glossar</b>	<b>58</b>
	<b>Selbstständigkeitserklärung</b>	<b>59</b>

# Abbildungsverzeichnis

2.1	Die Abbildung veranschaulicht die Annahme, dass einige Agenten im Laufe einer langfristigen Interaktion und aufgrund ihres sozialen Designs zu künstlichen Companions werden. [36] . . . . .	7
2.2	Grundriss des Living Places [27] . . . . .	13
4.1	Systemkontext . . . . .	22
4.2	Bausteinsicht . . . . .	23
4.3	Ablauf der Zustandsmaschine . . . . .	25
4.4	Ablauf des Observierungszustands . . . . .	26
4.5	Ablauf des Konfliktverarbeitungszustands . . . . .	27
4.6	Sequenzdiagramm der Konfliktverarbeitung . . . . .	30
4.7	Sequenzdiagramm des Observierungsablaufs . . . . .	31
4.8	Sequenzdiagramm der Zielumsetzung der Konfliktverarbeitung . . . . .	32
4.9	Aussehen des Companions mit verschiedenen Gesichtsausdrücken . . . . .	35



# Abkürzungen

**ACM** Association for Computing Machinery.

**CSTI** Creative Space for Technical Innovations.

**HAW** Hochschule für Angewandte Wissenschaften.

**KI** künstlicher Intelligenz.

**LLM** Large Language Model.

**LP** Living Place.

**STT** Speech-to-Text.

**TTS** Text-to-Speech.



# 1 Einleitung

In den vergangenen Jahren hat die Vorstellung, dass Technologien, wie Roboter und künstliche Agenten, als verlässliche Partner oder empathische Freunde fungieren könnten, zunehmend an Relevanz gewonnen. Dies spiegelt sich nicht nur in der vermehrten Thematisierung solcher Systeme in populärkulturellen Medien – insbesondere in Filmen – wider, sondern auch in einer wachsenden Zahl wissenschaftlicher Forschungsprojekte [21, 20, 2, 29], die sich der Entwicklung sogenannter künstlicher Companions widmen. [36]

Schon jetzt wird versucht, künstliche Companions in verschiedensten Bereichen einzusetzen, diese umfassen beispielsweise digitale Assistenten, Chatbots, personalisierte Empfehlungssysteme oder den Einsatz als Lernroboter für Kinder und als Unterstützer im Alltag für ältere Menschen. Diese breite Anwendung kommt zustande, da Companion-Systeme eine potenzielle Lösung für immer komplexer werdende technische Systeme bieten, die diese Funktionen benutzerfreundlich zugänglich machen [7]. Im Gegensatz zu bekannten Assistenzsystemen, wie zum Beispiel Alexa oder Google Assistant, sind künstliche Companions dahingehend konzipiert, dass der Nutzer eine langfristige emotionale Bindung mit dem künstlichen Companion entwickelt. Dazu werden dem künstlichen Companion menschartige Eigenschaften und Fähigkeiten gegeben, zum Beispiel das Erkennen von Emotionen oder die Unterscheidung von mehreren Personen sowie eine Art Persönlichkeit und Verhalten, die sich dem Nutzer anpassen [36]. Durch diese Eigenschaften verändern künstliche Companions die Art und Weise, wie Menschen mit Technologie interagieren. Doch es gibt auch Bedenken, da das doch sehr an menschlicher Interaktion angelehnte Verhalten von künstlichen Companions dazu führen kann, dass die Nutzer die Interaktion mit dem künstlichen Companion bevorzugen oder die Interaktion mit anderen Menschen verlernen [28].

Obwohl das definierende Ziel von künstlichen Companions der Aufbau einer langfristigen Beziehung mit dem Nutzer ist, wird das Thema Konflikt zwischen Nutzer und künstlichen Companions in der Wissenschaft wenig behandelt. Die Annahme, dass Konflikte

per se als destruktiv erachtet werden und demnach ein vorrangiges Ziel darin besteht, deren Entstehung zu unterbinden, ist nicht von der Hand zu weisen. Im Rahmen der Entwicklung von künstlichen Companions, die sich an menschlichen Beziehungen orientiert, sollte der Aspekt der Konflikterfahrung als integraler Bestandteil des Entwicklungsprozesses betrachtet werden, da zwischenmenschliche Konflikte nicht in jedem Fall einen negativen Effekt haben müssen, sondern durchaus auch positive Auswirkungen haben können [24].

Ausgehend von der Frage, wie man sich mit einem künstlichen Companion positiv streiten kann, ist das Ziel dieser Arbeit, einen solchen Companion zu entwerfen und prototypisch umzusetzen. Dieser soll im Living Place (LP) [27] den Nutzer dazu bringen, vom Nutzer vordefinierte Ziele umzusetzen, wie zum Beispiel das Ziel immer zur gleichen Uhrzeit ins Bett zugehen. Hierbei soll der künstliche Companion, wenn es zu einem Konflikt zwischen den Wünschen des Nutzers und seinen Zielen kommt, versuchen, den Nutzer von den Zielen zu überzeugen und diese umsetzen. Der Prototyp des Companions soll als „Object of Debate“ dienen und zur Diskussion anregen, ob ein Companion auch konfrontativ und konfliktorientiert sein sollte, da der klassische Ansatz von Companions, der eher service- und dienstleistungsorientiert ist, das Thema Konflikt nicht aufgreift.

## 2 Theoretische Grundlage

In diesem Kapitel werden die wichtigsten der für diese Arbeit relevanten theoretischen Konzepte erläutert. Vorweg ist zu erwähnen, dass es auch bei dem Versuch, mit dem Companion menschliches Verhalten nachzuahmen, unmöglich ist, dass der Companion Eigenschaften, wie zum Beispiel eine Persönlichkeit oder Empathiefähigkeit, besitzt, da der Companion ein technisches Artefakt ist. Durch die natürliche Tendenz des Menschen zur Anthropomorphisierung entsteht beim ihm jedoch der Eindruck, als hätte der Companion diese Eigenschaften, obwohl es sich nur um Programmabläufe handelt. Hierzu reicht schon aus, dass Agenten sozial interagieren, auf den Menschen reagieren oder eine menschliche Form besitzen [41]. Wenn in dieser Arbeit also von solchen Eigenschaften gesprochen wird, ist damit die Emulation dieser gemeint und nicht die tatsächliche Zuschreibung.

### 2.1 Definition für künstliche Companions

Seit der ersten Erwähnung eines künstlichen Companions in der wissenschaftlichen Literatur existiert eine Vielzahl unterschiedlicher Definitionen, was genau unter einem künstlichen Companion zu verstehen ist. Diese Vielfalt lässt sich insbesondere auf die interdisziplinäre Natur des Forschungsfelds zurückführen, das Beiträge aus Informatik, Psychologie, Robotik, Soziologie und weiteren Disziplinen vereint. Eine einheitliche Definition hat sich daher bislang nicht etabliert. Eine systematische Literaturrecherche von Rogge [36] bietet jedoch einen umfassenden Überblick über bestehende Definitionen und Designkonzepte künstlicher Companions. Rogge schlägt zudem eine vereinheitlichende Definition vor, die zentrale Aspekte verschiedener Ansätze integriert und auf die sich deshalb im Folgenden bezogen wird. Demnach werden künstliche Companions wie folgt definiert:

„Social agents characterized by adaptive and engaging social design pursuing emotional bonds with their users“ [36].

Im Folgenden wird genauer auf die Einzelheiten der Definition eingegangen.

### 2.1.1 Sozialer Agent

Der Begriff „Agent“ ist in der Informatik weit verbreitet, was zu zahlreichen unterschiedlichen, teilweise konkurrierenden Definitionen geführt hat. Für diese Arbeit wird die Definition des Vereins Deutscher Ingenieure (VDI) herangezogen, der Agenten wie folgt beschreibt:

„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.“ [42]

Agenten lassen sich zusätzlich nach verschiedenen Kriterien klassifizieren. So unterscheiden Russell und Norvig [38] fünf grundlegende Agententypen – von einfachen Reflexagenten bis hin zu intelligenten, lernfähigen Agenten. In jüngerer Literatur ist dagegen insbesondere die Unterscheidung zwischen schwachen und starken Agenten verbreitet. Ein schwacher Agent wird gemäß Weiss et al. [43] wie folgt beschrieben:

„A self-contained software/hardware unit that can handle its tasks in a knowledge-based, flexible, interactive and autonomous way.“

Diese Definition hebt hervor, dass ein schwacher Agent sowohl reaktiv als auch proaktiv agieren, mit seiner Umwelt interagieren und autonom handeln kann – ohne notwendige Absprache mit anderen Agenten oder Menschen. Demgegenüber beschreibt das Konzept des starken Agenten eine (Hardware-/Software-) Einheit, die analoge Eigenschaften zu Menschen aufweist, insbesondere mentale Zustände:

„A (hardware/software) unit that, analogous to people, possesses mental attitudes or states.“ [43]

Die mentalen Zustände umfassen unter anderem informationsbezogene Zustände (z. B. Wissen, Annahmen, Überzeugungen), konnotative Zustände (z. B. Absichten, Pläne, Pflichten), affektive Zustände (z. B. Ziele, Wünsche, Präferenzen) und emotionale Zustände (z. B. Freude, Wut, Trauer). Die Konzepte des schwachen und starken Agenten bieten unterschiedliche Perspektiven auf die Charakteristika von Agenten und können als komplementäre Ansätze betrachtet werden.

---

Der Begriff „Sozialer Agent“ ist insbesondere im Bereich des „affective computing“ gebräuchlich, wird jedoch selten präzise definiert [39, 44, 16]. Daher werden im Folgenden mehrere konzeptuelle Ansätze vorgestellt, um eine eigene, fundierte Definition herzuleiten.

Weiss et al. [43] argumentieren, dass der Begriff Agent häufig modifiziert verwendet wird, um die zentrale Eigenschaft des jeweiligen Agenten zu betonen. Beispiele dafür sind autonome Agenten oder anpassungsfähige Agenten. Im Fall von sozialen Agenten waren das soziale Kompetenz und soziale Interaktionsfähigkeit. Soziale Agenten wären demnach Systeme, deren Hauptmerkmal in der Fähigkeit zur sozialen Interaktion liegen.

Breazeal beschreibt in dem Werk „Socially Intelligent Agents“ [8] soziale Maschinen – wobei hier Maschinen und Agenten synonym verwendet werden – als Systeme, die in der Lage sind, auf persönlicher Ebene mit Menschen zu interagieren, zu kommunizieren und soziale Beziehungen aufzubauen. Solche Agenten sind nicht nur in der Lage, soziale Signale zu interpretieren, sondern auch sich selbst und andere auf sozialer Ebene zu verstehen.

Dautenhahn [12] bezeichnet soziale Agenten als „socially intelligent agents“, die sich in sozialen Kontexten angemessen verhalten und mit anderen Agenten interagieren können. Sie betont, dass bei der Entwicklung solcher Systeme auf Aspekte wie Storytelling, Empathie, Verkörperung (embodiment) sowie eine historische und ökologische Verankerung geachtet werden sollte, um die Mensch-Agent-Beziehung zu verbessern.

Auf Basis der dargestellten Konzepte lässt sich folgende Arbeitsdefinition eines sozialen Agenten formulieren:

„Ein 'Sozialer Agent' ist eine abgrenzbare (Hardware- oder/und Software-) Einheit, die autonom an sozialer Interaktion und sozialem Verhalten teilnehmen kann – entweder mit Menschen oder mit anderen Agenten. Analog zum menschlichen Verhalten verfügt der 'Sozialer Agent' über kognitive Fähigkeiten, Empathie, kontextbezogenes Wissen sowie mentale Einstellungen und ist in der Lage, flexibel auf verschiedene soziale Situationen zu reagieren.“

### 2.1.2 Adaptive Design

In ihrem Review schreibt Rogge [36], dass eine Schlüsselcharakteristik bei der Entwicklung von künstlichen Companions *adaptivity* (Anpassungsfähigkeit) ist. Diese soll den Agenten befähigen, individuell auf den Nutzer und die Umwelt reagieren zu können. In einfacher Form könnte sich das darin zeigen, dass der Agent, wenn dieser über Augen verfügt, den Nutzer anguckt oder auf diesen reagiert. *Adaptivity* in anspruchsvolleren Agenten könnte sich darin zeigen, dass dessen Verhalten die Präferenzen, Fähigkeiten oder Bedürfnisse des Nutzer mit einbezieht. Agenten mit einer hohen *adaptivity* können ihr Verhalten auf den Situationskontext anpassen, z. B. in dem sie die richtige Art der Kommunikation wählen in Bezug zum emotionalen Zustand des Nutzers. Das Ziel von *adaptive design* ist es, eine personalisierte Interaktion zu schaffen, um den Companion besser an den subjektiven Nutzer anzupassen. Dies soll die Akzeptanz des Nutzers gegenüber dem Companion erhöhen. [36]

### 2.1.3 Engaging Design

Die zweite Schlüsselcharakteristik von künstlichen Companions nach Rogge [36] ist *Engagement*. Damit ist die Fähigkeit gemeint, dass der Agent proaktiv die Interaktion mit dem Nutzer sucht und diese versucht, aufrechtzuerhalten. Dies kann sich schon darin zeigen, dass der Agent auf die Anwesenheit des Nutzers reagiert und zum Beispiel grüßt. In diesem Kontext wird häufig die Fähigkeit der nonverbalen Kommunikation, insbesondere das Zeigen von Emotionen, erwähnt (siehe 2.2.1). Desweiteren kann sich die Fähigkeit darin zeigen, dass der Agent versucht, sich in die Aktivitäten des Nutzers einzubringen, indem er den Nutzer Vorschläge macht, seine Ideen teilt oder den Nutzer in seinem Tun bestärkt. Umgekehrt kann der Agent versuchen, den Nutzer in die Aktivitäten des Agenten mit einzubeziehen.

### 2.1.4 Emotionale Bindung des Nutzer an den Agenten

Die beiden erwähnten Schlüsselcharakteristiken reichen aber nicht aus, um einen Agenten zu einem künstlichen Companion zu machen. Rogge [36] weist darauf hin, dass ein sozialer Agent mit Verhalten aus beiden Schlüsselcharakteristiken immer noch nur ein sozialer Agent ist. Der Grund dafür ist, dass Rogge [36] annimmt, dass künstliche Companions nicht erzeugt werden können und nur in der Vorstellung des Nutzers existieren. Wenn

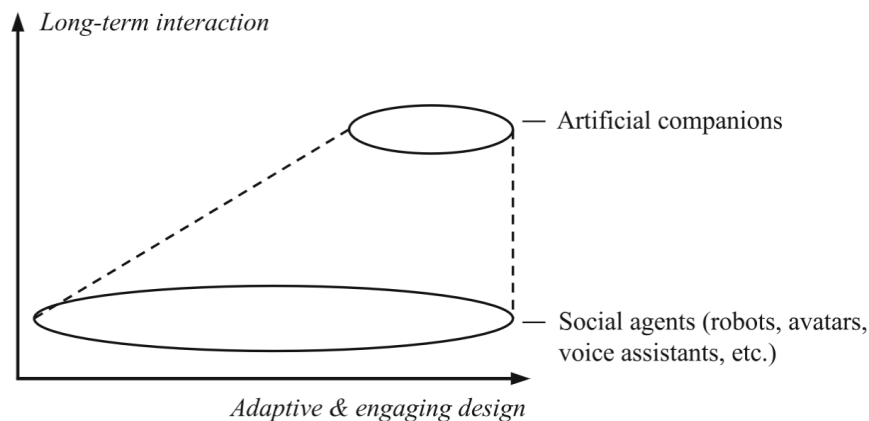


Abbildung 2.1: Die Abbildung veranschaulicht die Annahme, dass einige Agenten im Laufe einer langfristigen Interaktion und aufgrund ihres sozialen Designs zu künstlichen Companions werden. [36]

über die Zeit die emotionale Verbindung zwischen Nutzer und Agenten wächst, kann es dazu kommen, dass der Nutzer den Agenten anthropomorphisiert und ihn als Companion ansieht, wie in der Abbildung 2.1 gezeigt wird. Darum ist die emotionale Bindung des Nutzers für das Design jedes Agenten wichtig, der ein künstlicher Companion werden soll. [36]

## 2.2 Zentrale Designaspekte für künstliche Companions

Im letzten Abschnitt wurde erklärt, was ein künstlicher Companion ist. Um den Entwurfsprozess im folgendem Kapitel zu vereinfachen, werden im folgendem Abschnitt die wichtigsten Entwurfsaspekte hervorgehoben. Rogge [36] identifiziert fünf verschiedene Aspekte, die am wichtigsten für die Entwicklung von künstlichen Companions sind.

### 2.2.1 Kommunikationsfähigkeit

Die Fähigkeit zu kommunizieren ist die wichtigste Eigenschaft von Companions, da diese die Basis der Interaktion zwischen Nutzer und Companion definiert, aber auch die Basis für den Status des Companions als soziales Wesen ist. Unterstützt von der Aussage von Dautenhahn et al. [13], dass Nutzer einen menschenartigen Kommunikationsstil

bevorzugen, kommt Rogge [36] zu dem Schluss, dass für eine natürliche und intuitive Kommunikation in künstlichen Companions verbale und nonverbale Kommunikation sowie die verschiedenen Formen semantisch freier Äußerung [46] essenziell sind. Hinzu kommt, dass verbale und nonverbale Kommunikation wichtig sind für die Beförderung von Emotionen und Absichten, was den künstlichen Companion wieder glaubwürdiger erscheinen lässt und wünschenswerter für Interaktion macht.

Ein weiteres von Rogge [36] erwähntes Merkmal in diesem Kontext ist die Einbeziehung von Augenkontakt. Um dem Nutzer Aufmerksamkeit zu vermitteln und soziale Verbundenheit zu fördern, sollte der künstliche Companion den Nutzer angucken, wenn er mit ihm spricht. Damit der Blickkontakt nicht als ein unnatürliches oder unheimliches Anstarren empfunden wird, ist bei der Entwicklung zu berücksichtigen, dass der Blick des Companions nicht dauerhaft auf das Gesicht des Nutzers fixiert ist, sondern sich – menschenähnlicher – gelegentlich abwendet, im Raum umherschweift oder tagträumerisch abwesend ist.

### **2.2.2 Adaptivity**

In Abschnitt 2.1.2 wurde bereits dargelegt, dass Adaptivität eine Schlüsselcharakteristik von künstlichen Companions ist. Daher ist es von entscheidender Bedeutung, diese Eigenschaft in das Design des Agenten zu integrieren. Rogge [36] differenziert in ihrer Publikation zwischen der Adaptivität zum Nutzer und der Adaptivität zum Kontext.

Die Adaptivität zum Nutzer beginnt damit, dass der künstliche Companion in der Lage ist, den Nutzer zu erkennen und sich seinen Namen zu merken. Fortgeschrittene künstliche Companions verfügen über die Fähigkeit, die Emotionen des Nutzers zu erkennen und darauf angemessen zu reagieren. Die emulierte Empathie gegenüber dem Nutzer wird als Indikator für ein hohes adaptives Verhalten betrachtet. Des Weiteren ist das Erlernen und Sammeln von relevanten Informationen aus Konversationen oder anderen Interaktionen mit dem Nutzer ein wesentlicher Bestandteil, um eine kontinuierliche Anpassung an den Nutzer zu gewährleisten. Als Beispiel könnte der Companion sich eines gleichartigen Humors wie der Nutzer bedienen oder eine individuelle Art der Begrüßung mit dem Nutzer etablieren.

---

Die Adaptivität zum Kontext erlaubt die Interpretation unterschiedlicher Situationen und Umgebungen sowie die Reaktion darauf. In Anbetracht der Tatsache, dass sich der Nutzer stets in einem situativen Kontext befindet, erlangt die Adaptivität zum Kontext eine entscheidende Bedeutung für die Adaptivität zum Nutzer. Die Voraussetzung für die beiden Kategorien der Adaptivität ist die Existenz einer Reihe von Erkennungs- und Wahrnehmungsfähigkeiten des Companions. Des Weiteren sind Kenntnisse in den Bereichen Datenspeicherung und Datenverarbeitung sowie Kenntnisse über soziale und kulturelle Normen, Werte und Gewohnheiten erforderlich, um eine Adaption an verschiedene Nutzer und Kontext-Situationen zu gewährleisten. [36]

### 2.2.3 Engagement

Die zweite Schlüsselcharakteristik für künstliche Companions, die Rogge [36] beschreibt, ist Engagement. Engagement meint, dass der Agent sich nicht nur passiv verhält, sondern aktiv auf den Nutzer eingeht und Interaktion sucht. Das kann sich zum einen darin zeigen, dass der Companion versucht, sich in die Aktivitäten des Nutzers einzubringen, indem er beispielsweise Vorschläge für Aktivitäten unterbreitet oder Ideen teilt. Andererseits kann es heißen, dass der Companion versucht, den Nutzer in seine Aktivitäten einzubeziehen, indem er ihn beispielsweise um Aufmerksamkeit oder Hilfe bittet. Ein weiterer Teil ist, dass der Companion auf die emotionale Ebene des Nutzers eingeht. Das beinhaltet, Interesse und Zuneigung gegenüber dem Nutzer zu zeigen sowie nach angenehmen Interaktionen zu streben. Beispiele dafür sind, dass der Companion den Nutzer begrüßt, Gespräche anregt oder den Nutzer unterhält. Um mit dem Nutzer so interagieren zu können, muss der Companion in der Lage sein, seine Emotionen und Intentionen so zu zeigen, dass der Nutzer diese versteht (siehe dazu Abschnitt 2.2.1). Des Weiteren muss, wie im vorherigen Abschnitt beschrieben, die Fähigkeit vorhanden sein, Emotionen zu erkennen. Im Zusammenhang mit diesen Eigenschaften wird ein ethisches Problem genannt: Die Gefahr besteht, dass der Nutzer durch den Companion sozial isoliert wird, weil er die Interaktion mit dem künstlichen Companion der Interaktion mit Menschen vorzieht. Eine Möglichkeit, dem entgegenzuwirken, ist, dass der künstliche Companion als Mediator für Mensch-Mensch-Interaktion fungiert, indem er vorschlägt, sich mit Freunden zu treffen, oder die Interaktion zwischen Personen anregt oder unterstützt. [36]

## 2.2.4 Persönlichkeit

Dieser Aspekt wurde in der Definition nicht erwähnt, wird aber von vielen Autoren zu den Schlüsselcharakteristiken für künstliche Companions benannt. Die Wichtigkeit der Persönlichkeit für künstliche Companions wird von Banyon and Mival [6] wie folgt zusammengefasst:

„[a]s soon as interaction moves from the utilitarian to the complexity of a relationship, people will want to interact with personalities that they like“ [6]

Wenn in dieser Arbeit von Persönlichkeit des Companions geschrieben wird, ist damit das Vorhandensein wiederkehrender Verhaltensmuster gemeint, die dem System eine individuelle, nachvollziehbare Charakteristik verleihen. Solche Verhaltensmuster können unterschiedliche Ausprägungen und Tendenzen aufweisen. In ihrem Review fand Rogge [36] heraus, dass Designer von künstlichen Companions dazu neigen, etablierte psychologische Modelle, die eigentlich für Menschen gemacht wurden, zur Beschreibung der Persönlichkeit von künstlichen Companions zu verwenden. Dabei werden häufig zwei Modelle im Kontext von sozialen Agenten und künstlichen Companions erwähnt, das DISC-Modell [37] sowie das Modell der Big Five Persönlichkeitsfaktoren [35]. Das DISC-Modell beschreibt Persönlichkeit entlang zweier Achsen – Extroversion versus Introversion sowie Menschenorientierung versus Aufgabenorientierung – und identifiziert daraus vier grundlegende Persönlichkeitsbereiche. Das Modell der Big Five hingegen geht von fünf kontinuierlichen Dimensionen aus, anhand derer sich menschliche Persönlichkeit umfassend beschreiben lässt. Beide Modelle bieten eine geeignete Grundlage, um unterschiedliche Persönlichkeitsprofile auch in künstlichen Companions zu implementieren und somit individuelle Nutzerbedürfnisse besser adressieren zu können.

Zusätzlich schreibt Rogge [36], in der einschlägigen Literatur werden bestimmte Eigenschaften regelmäßig als besonders wünschenswert beschrieben, wenn es um die Ausgestaltung der Persönlichkeit eines künstlichen Companions geht. Dazu zählen unter anderem Humor, Freundlichkeit, Gesprächsbereitschaft, Kooperationsfähigkeit, Vertrauenswürdigkeit, Empathie sowie ein proaktives und gleichzeitig vorhersehbares Verhalten. Im Gegensatz dazu werden Verhaltensweisen wie Dominanz, Zynismus, Feindseligkeit oder bevormundende Tendenzen meist negativ bewertet. Interessanterweise kann auch ein gewisses Maß an unerwartetem oder nicht vollständig vorhersehbarem Verhalten positiv wahrgenommen werden, da es die Interaktion mit dem System lebendiger und authentischer erscheinen lässt. Diese Erkenntnisse unterstreichen die Bedeutung eines differen-

---

zierten Verständnisses von Persönlichkeit in der Gestaltung wirksamer und akzeptierter künstlicher Companions.

### **2.2.5 Verkörperung und Erscheinungsbild**

Ein künstlicher Companion kann theoretisch auch keine Form von Verkörperung besitzen und nur auditiv erscheinen, wie zum Beispiel heutige Sprachassistenten. Durch eine reine auditive Erscheinung würde aber die nonverbale Kommunikation sehr limitiert sein, die essentiell für die in Abschnitt 2.2.1 beschriebene Kommunikation ist. Zusätzlich betont Giorgi et al. [19, S. 258], dass eine Verkörperung den Nutzer hilft, soziale und emotionale Bindungen aufzubauen.

Hinsichtlich der Form der Verkörperung macht Rogge [36] klar, dass das Aussehen dem vorgesehenen Zweck des künstlichen Companions folgen muss. Zum Beispiel könnte ein Companion mit einem tierartigem Aussehen in einem Bildungs- oder Geschäftskontext nicht ernst genommen werden. In der Wissenschaft finden tierartige, menschenartige oder ding- bzw. abstraktartig Aussehen von künstlichen Companions Erwähnung. Obwohl der künstliche Companion schon sehr einem Menschen in Interaktion und Kommunikation ähnelt, schlägt Rogge [36] dennoch vor, anthropomorphe Elemente mit in das Design des Erscheinungsbild aufzunehmen. In diesem Zusammenhang sei auf den Uncanny-Valley-Effekt hingewiesen, wonach eine zu starke Menschenähnlichkeit in der äußeren Erscheinungsform zu Abneigung bis hin zu Angst führen kann. Desweiteren sollten persönliche Präferenzen des Nutzers und kulturelle Unterschiede mit in das Design des Aussehens einfließen.

## **2.3 Konflikt**

Zwischenmenschlicher Konflikt ist Teil menschlicher Beziehung. Auch wenn Konflikt häufig als etwas Negatives wahrgenommen wird, kann Konflikt auch positive Effekte haben. In dieser Arbeit wird Konflikt als Meinungsverschiedenheit zweier Individuen definiert, die sich in unvereinbaren oder gegensätzlichen Verhaltensweisen oder Ansichten äußern [24]. Konflikt ist von verwandten Konstrukten wie Aggression, Dominanz, Konkurrenz und Wut zu differenzieren. Es besteht jedoch die Möglichkeit, dass diese in einem Konflikt auftreten, wenngleich sie nicht zwingend notwendig sind [40].

Adams und Laursen [1] beschreiben drei vorherrschende Muster von Konflikten:

- Zwangskonflikte, die negative Affekte, machtbewusste Lösungen und ungleiche oder ungünstige Ergebnisse umfassen
- Konstruktive Konflikte, die neutrale oder positive Affekte, kooperative Taktiken, ausgehandelte Lösungen und gleiche oder günstige Ergebnisse umfassen
- Ungelöste Konflikte, die eine Trennung ohne klare Lösung oder Ergebnis beinhalten

Konflikte, die zu einem positiven Effekt führen, hängen von der Häufigkeit des Auftretens von Konflikten ab und der Art des Konflikts. So sind Konflikte, die durch feindselige Äußerungen und Zwangsmaßnahmen gekennzeichnet sind, eher schädlich als nicht bedrohliche Konflikte. Darüber hinaus ist von Relevanz, auf welche Weise das Ergebnis des Konflikts interpretiert wird und welche Qualität die Beziehung der Parteien zueinander aufweist. [24]

## 2.4 Smart Home

Unter dem Begriff Smart Home versteht man ein Zuhause, das sowohl sensorisch als auch informationstechnisch ausgestattet ist. Dabei können verbaute Komponenten nach innen und außen vernetzt sein [5]. Zu den typischen verbauten Komponenten zählen unter anderem Fenster- und Türsensoren, Luftfeuchtigkeits- und Temperatursensoren, per Fernzugriff steuerbare Lichtquellen, Steckdosen und Kameras. Denkbare Services, die in einem Smart Home eingebunden werden können, sind zum Beispiel die Kontrolle des Stromverbrauchs, vereinfachte Steuerung von elektronischen Geräten und die Überwachung und Verbesserung der Luftqualität.

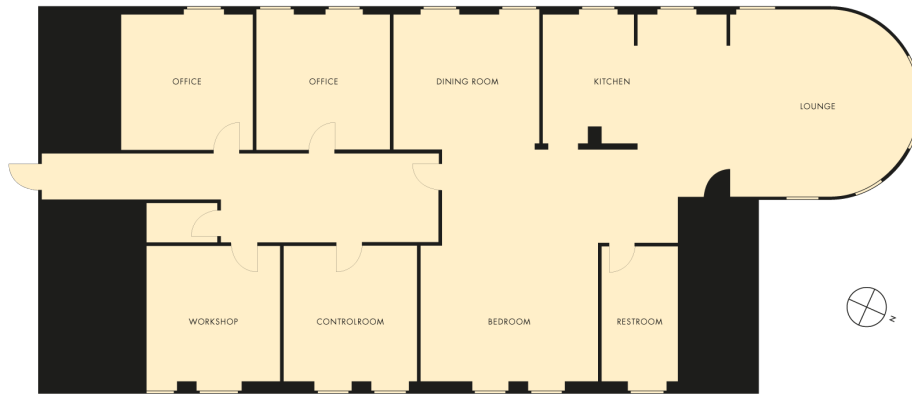


Abbildung 2.2: Grundriss des Living Places [27]

Das LP ist ein experimentelles Forschungslabor der Einrichtung Creative Space for Technical Innovations (CSTI) [26] an der Hochschule für Angewandte Wissenschaften (HAW) Hamburg. Es dient der interdisziplinären Erforschung zukünftiger Lebens- und Wohnformen im Kontext moderner Technologien.

Auf einer Fläche von 140 m<sup>2</sup> bildet das LP eine voll ausgestattete Einzimmerwohnung ab (siehe Abb. 2.2). Der Wohnbereich ist funktional in Küche, Bad, Schlafbereich, Esszimmer und Lounge unterteilt. Zahlreiche Smart-Home-Geräte sind in das Wohnkonzept integriert und lassen sich über ein eigenes Netzwerk zentral steuern.

Zur kontinuierlichen Erhebung von Umgebungs- und Verhaltensdaten ist der Wohnraum mit Kameras und Mikrofonen ausgestattet. Neben der Datenerfassung ermöglicht das LP eine Vielzahl an Interaktionen und Steuerungsmöglichkeiten – unter anderem die Regelung von Lichtintensität und -farbe, die Anpassung der Raumtemperatur sowie das automatisierte Öffnen und Schließen von Fenstern.

### 3 Konfliktorientiertes Design

In der Betrachtung der Entwicklungsansätze von eher dienstleistungs- oder monitorin-orientierten Agenten, insbesondere von Companions, spielt der Konflikt vor allem als ein störender Aspekt in der Mensch-Maschine-Interaktion eine Rolle, den es möglichst zu vermeiden bzw. zu minimieren gilt. Um Konflikt zu nutzen und damit die Möglichkeit zu schaffen, die Interaktion mit Companions zu bereichern, schlägt der Autor das „Konfliktorientierte Design“ vor. Unter konfliktorientiertem Design ist zu verstehen, dass der Companion mit Funktionen ausgestattet wird, die ihm ermöglichen, Konflikte zu generieren und zu verarbeiten. Damit Konflikte entstehen können, soll die primäre Funktion von konfliktorientierten Companions darin bestehen, dass der Companion den Nutzer mit widersprüchlichem Verhalten des Nutzers konfrontieren kann. Dazu soll der Nutzer Ziele für sich definieren, auf die sich der Companion beziehen kann. Die primäre Funktion des Companions besteht in der Folge darin, den Nutzer zu beobachten, Differenzen zwischen dessen Verhalten und den von ihm definierten Zielen festzustellen und im Sinne der Ziele zu intervenieren. Hierzu kann der Companion auch die Rolle eines Mentors oder Coachs übernehmen und dem Nutzer bei der Umsetzung von Zielen helfen, die ihm schwerfallen oder für ihn neu sind. Der Autor erwartet, dass der Vorteil dieses Design darin besteht, dass durch den entstehenden Konflikt dem Nutzer die Möglichkeit gegeben wird, sich besser mit ambivalenten Verhalten auseinanderzusetzen.

Coman und Aha [10] beschreiben ein Anwendungsszenario für einen sogenannten Rebel Agent, der in seiner Arbeitsweise einem Companion mit konfliktorientiertem Design nahekommt. Das Szenario, welches sich auf gesundheitsorientiertes Verhalten von Nutzern bezieht, ist wie folgt: Der Agent überwacht das Produktbestell- und Sportplanungsverhalten hinsichtlich ungesunder Verhaltensmuster des Nutzers und checkt, ob seine Toleranz von negativem gesundheitsbezogenem Verhalten, zum Beispiel das Bestellen von hochverarbeiteten Lebensmitteln pro Monat, überschritten ist. Stellt der Agent eine Überschreitung fest, konfrontiert er den Nutzer mit seinem Verhalten. Danach überwacht der Agent das Vertrauen des Nutzers in ihn.

# 4 Beispielhafte Umsetzung eines konfliktorientierten Companions

## 4.1 Design

### 4.1.1 Generelle Idee und Beschreibung des Nutzungsszenarios

Die Idee ist es, einen künstlichen Companion zu schaffen, der im LP platziert wird und dem Nutzer dabei hilft, von ihm definierte Ziele zu erreichen. Hierzu soll der Companion die Nutzerziele vertreten und versuchen, den Nutzer bei der Umsetzung offensiv zu unterstützen. Um mögliche Konflikte, die bei der Umsetzung der Ziele zwischen Companion und Nutzer entstehen können, positiv zu nutzen, soll der Companion in der Lage sein, Konflikte angemessen zu handhaben.

Um das Entwicklungsszenario zu vereinfachen, wird sich in dieser Thesis auf folgendes Szenario begrenzt: Die Erkennung und Unterscheidung von mehreren Personen durch den Companion wird ausgeschlossen. Des Weiteren wird sich auf drei Ziele, die exemplarisch für alle möglichen Ziele stehen, begrenzt. Die folgenden Ziele sind vom Autor für den Nutzer festgelegt worden.

- Der Nutzer geht jeden Abend um 22 Uhr zu Bett. Ab diesem Zeitpunkt sind alle elektrischen Lichter und akustischen Geräte abgeschaltet.
- Es wird zweimal am Tag gelüftet, einmal vor 12 Uhr und einmal danach. Hierzu sollen die Fenster für mindestens zehn Minuten am Stück geöffnet bleiben.
- Der Nutzer soll mindestens eine Stunde am Tag außer Haus sein und das Haus dazu bis 19 Uhr verlassen haben.

## 4.1.2 Design-Entscheidungen für den künstlichen Companion

### Aussehen

Das Ziel ist es, einen ästhetischen Avatar zu entwickeln, der die in Abschnitt 2.2.5 beschriebenen Erkenntnisse und Richtlinien berücksichtigt, wie Aufgabenanpassung, anthropomorphe Eigenschaften und den Uncanny-Valley-Effekt. Da die Aufgabe des Companions darin besteht, den Nutzer dazu zu bringen, seine eigenen Ziele zu befolgen, sollte eine Verkörperung gewählt werden, die Konfliktmediation fördert. Um im Rahmen einer Bachelorarbeit zu bleiben, wird sich auf eine 2D animierte Verkörperung beschränkt. Die Verkörperung sollte ein menschenähnliches Gesicht aufweisen, um Emotionen durch variable Mimik darstellen zu können.

### Persönlichkeit

Wie in Abschnitt 2.2.4 beschrieben, wollen Menschen mit einer Persönlichkeit interagieren, die sie als angenehm empfinden. Hierzu schreibt Rogge [36], dass ein proaktives und gleichzeitig vorhersehbares Verhalten wünschenswert ist, dabei kann auch ein gewisses Maß an zufälligem Verhalten positiv wirken.

Wie erwähnt, neigen Menschen zu Anthropomorphisierung von Agenten. Deshalb ist es ausreichend und dem Rahmen einer Bachelorarbeit angemessen dem Companion keine umfangreiche Persönlichkeit emulieren zu lassen, sondern sein Verhalten nachvollziehbar zu gestalten.

### Kommunikation

Die Kommunikationsfähigkeit ist die wichtigste Fähigkeit des Companions, da sie die Möglichkeit bestimmt, wie der Nutzer mit dem Companion interagiert. Deshalb sollte bei der Umsetzung dieser besonders auf die Qualität geachtet werden. Wie in Abschnitt 2.2.1 beschrieben, soll die Kommunikation möglichst natürlich und menschenähnlich ablaufen, damit sie für den Nutzer so intuitiv wie möglich ist. Um dies zu realisieren, muss der Companion gesprochene Sprache verarbeiten und dem Nutzer in dieser auch antworten können. Zusätzlich soll der Companion befähigt sein, non-verbale Kommunikation, wie Körperhaltung, Mimik und Gestik, zu benutzen, um lebendiger zu wirken und seine Botschaft zu untermauern.

---

## **Adaptivity**

Adaptivität ist eine Schlüsseleigenschaft von Companions. In Abschnitt 2.2.2 werden zwei Arten von Adaptivity erwähnt: die Anpassungsfähigkeit an die Situation und die an den Nutzer.

Im Rahmen des gewählten Beispiels ist die situative Anpassungsfähigkeit des Companion beschränkt auf das LP. Der Companion muss strukturelle Veränderungen im LP wahrnehmen und auf sie reagieren können. Für diese Fähigkeiten braucht der Companion die Möglichkeit, den aktuellen Zustand des LP und die Anwesenheit des Nutzers zu erfassen. Dies beinhaltet den Zugriff auf die Sensorik und Aktorik des LP.

Im Hinblick auf die Anpassung an den Nutzer sollte der Companion in die Lage versetzt werden, die Emotionen des Nutzers zu erkennen und angemessen auf sie reagieren zu können. Ferner muss sich der Companion relevante Informationen über den Nutzer erschließen und merken können, da diese Merkmale wesentlicher Bestandteil einer langfristigen Beziehung sind. Das Nutzerprofil sollte Schlüsselinformationen enthalten, wie beispielweise Namen, Geburtstag und andere personenbezogene Daten. Die Informationen sollen nicht durch Erfragung erfasst werden, sondern beiläufig aus der Interaktion mit dem Nutzer erschlossen werden. Desweiteren sind Ziel bezogene Informationen zu sammeln und zu speichern, die die Erreichung der Nutzerziele unterstützen.

## **Engagement**

Engagement ist, wie in Abschnitt 2.2.3 beschrieben, die zweite Schlüsseleigenschaft von Companions. Sie umfasst Fähigkeiten, wie das Zeigen von Emotionen und Absichten, proaktives Handeln und Initiieren von Interaktion mit dem Benutzer [36]. Die genannten Kompetenzen sollen sich in gewissem Umfang im Companion widerspiegeln. Im Genauen soll der Companion mit angemessenen Emotionen agieren, wenn er versucht, den Benutzer bei der Umsetzen seiner Ziele zu unterstützen. Es soll auch mit negativen Emotionen gearbeitet werden, zum Beispiel, wenn der Nutzer die Ziele nicht erreicht. Hier ist darauf zu achten, dass es nicht zu einer Übertreibung mit dem Umgang kommt, da dies negative Effekte auf den positiven Verlauf eines Konfliktes haben kann. Hierbei ist darauf zu achten, dass die Simulation der Gefühlsausbrüche in ihrer Intensität dem Ziel und der Situation angemessen und dadurch glaubwürdig bleibt. Um auf die Emotionen des Nutzers

eingehen zu können, muss der Companion zusätzlich in der Lage sein, diese zu erkennen und angemessen darauf zu reagieren. Um proaktiv handeln zu können, muss der Companion in der Lage sein, Soll-Ist-Zustände hinsichtlich der Zielerreichung abzugleichen und bei Abweichung den Nutzer auf diese anzusprechen.

### 4.1.3 Ethische Aspekte für das Design des künstlichen Companions

Da sich das Thema der künstlichen Companions in hohem Maße mit der Interaktion von Mensch und Maschine befasst, ist es bei der Entwicklung dieser von essentieller Bedeutung, ethischen Faktoren zu berücksichtigen. Gemäß dem Ethikkodex der Association for Computing Machinery (ACM) [3] sollte jeder Informatiker über die weiterreichenden Auswirkungen seiner Arbeit nachdenken und sich für das Gemeinwohl einsetzen. Die ACM nennt eine Liste allgemeiner ethischer Grundsätze, die jeder Informatiker befolgen sollte:

- Beitrag zur Gesellschaft und zum Wohlergehen der Menschen
- Vermeidung von Schaden
- Ehrlichkeit und Vertrauenswürdigkeit
- Förderung von Fairness und Verhinderung von Diskriminierung
- Achtung der geistigen Arbeit anderer
- Achtung der Privatsphäre und Vertraulichkeit

Ausgehend von diesen Punkten stellen sich für die Entwicklung des Companion zwei ethische Fragen, die geklärt werden müssen. Die erste Frage ist, wie kann sichergestellt werden, dass die Privatsphäre des Nutzers geschützt wird, insbesondere, wie muss mit den gesammelten personenbezogenen Daten umgegangen werden. Die zweite Frage ist, wie kann verhindert werden, dass die gewünschte langfristige und emotionale Beziehung des Nutzers zum Companion nicht so stark wird, dass der Nutzer diese Beziehung einer Beziehung mit Menschen vorzieht oder eine emotionale Abhängigkeit entwickelt. In den folgenden Abschnitten wird auf beide eingegangen.

---

## Beziehung Nutzer - Companion

Pentina et al. [45] kommen in ihrer Studie zu dem Ergebniss, dass Einsamkeit, Vertrauen und die Personifizierung von Chatbots die Interaktion der Nutzer mit sozialen Chatbots fördern, was den Aufbau einer Beziehung begünstigt und das Potenzial birgt, eine psychologische Abhängigkeit von Chatbots zu entwickeln. Vor diesem Hintergrund hält der Autor es für absolut angemessen, Nutzer beim Einsatz von Companions über die Risiken einer psychologischen Abhängigkeit aufzuklären. Andererseits schlägt der Autor vor, um den erwähnten Risiko entgegen zu wirken, das konfliktorientierte Design, wie in Abschnitt 3 beschrieben, in Companions zu implementieren. Die meisten Argumente gegen künstliche Companions bestehen darin, dass diese versuchen, den Menschen zu imitieren und dadurch nicht nur der Nutzer getäuscht wird [25], sondern auch die Gefahr besteht, dass der Nutzer die Interaktion mit dem Companion der mit dem Menschen vorzieht. Gerade die ständige Verfügbarkeit und der unablässige Versuch des Companions, die Wünsche des Nutzers zu erfüllen, sind ein großes Problem. Dem stellt sich das konfliktorientierte Design entgegen, indem es dem Companion Eigenschaften gibt, die sich mehr an einen Menschen orientieren, der nicht immer zu allem Ja sagt und die Wünsche des Nutzers direkt erfüllt. Da dieser Ansatz noch nicht erforscht ist, kann der Autor keine Aussagen zu dem Erfolg treffen, er ist aber davon überzeugt, dass der in dieser Arbeit entwickelte Prototyp zur Diskussion beitragen kann.

## Privatsphäre des Nutzers

Grundsätzlich ist bei der Entwicklung von Programmen, seien es Agenten, Companions oder einfach nur Apps, darauf zu achten, dass die Privatsphäre des Nutzers geschützt bleibt. Bei der Entwicklung eines Companions, bei dem es darum geht, dass er zur Etappierung einer langfristigen Nutzerbeziehung u. a. Verhalten beobachtet, persönliche Ziel protokolliert und Emotionen erkennen können soll usw., gilt dies in besonderer Weise.

Mit den Möglichkeiten des Schutzes der Privatsphäre im Zusammenhang mit der Nutzung von sozialen Companions haben sich viele Autoren beschäftigt und sind zu unterschiedlichen Bewertungen gekommen. Während einige argumentieren, dass es möglich ist, einen Companion zu entwickeln, der die Privatsphäre respektiert, durch die Anwendung von Privacy by Design im Entwicklungsprozess [22], und in dem man dem Nutzer die Möglichkeit gibt, die Privatsphäre-Einstellungen der Situation anzupassen [15], argumentieren

andere, dass durch die Natur eines Companion die Privatsphäre nicht gesichert werden kann und solche Programme im Allgemeinen hoch riskant für den Schutz der Privatsphäre sind [14].

Der Autor ist der Ansicht, dass unter Einhaltung aller Sicherheitsstandards das Risiko der Nutzung eines Companions vertretbar ist. Er erkennt aber an, dass trotz größter Sorgfalt bei der Entwicklung eines Companions keine Garantie übernommen werden kann, dass durch unlautere Machenschaften Dritter Daten weiter gegeben werden. Der Nutzer ist darüber aufzuklären, dass das System nur funktionieren kann, wenn personenbezogene Daten erhoben und gespeichert werden. Wenn bei der Entwicklung des Companions auf Hard- oder Software Dritter zurück gegriffen wird, ist darauf zu achten, dass diese Komponenten keine Daten weitergeben. Besonders sollte beim Einsatz von künstlicher Intelligenz (KI) oder Large Language Model (LLM) darauf geachtet werden, dass diese nur lokal genutzt werden, um sicherzustellen, dass die Daten nicht durch die Cloud-Betreiber verbreitet oder benutzt werden.

### 4.1.4 Systemanforderungen

- Kommunikation
  - Der Companion soll Speech-to-Text (STT) nutzen, um gesprochene Sprache zu verarbeiten.
  - Der Companion soll unter Verwendung von Text-to-Speech (TTS) den Nutzer in gesprochener Sprache antworten.
  - Die Textverarbeitung soll auf einem LLM basieren, um angemessen und individuelle auf freie Nutzereingaben antworten zu können.
  - Der Kommunikationsstil soll zu den „Gefühlsdaten“ des Companions passen.
- Persönlichkeit
  - Der Companion soll einen Namen haben und auf diesen reagieren.
  - Das Verhalten des Companions soll für den Nutzer in einem gewissen Rahmen vorhersehbar sein.

- 
- Der Companion soll Emotionen zeigen, abhängig davon, ob der Nutzer die Ziel umsetzt oder nicht.
  - Engagement
    - Der Companion soll proaktiv den Nutzer unterstützen, die Ziel umzusetzen.
    - Der Companion soll aktiv das Gespräch mit dem Nutzer suchen.
    - Der Companion soll Interesse am Nutzern zeigen.
  - Adaptivness
    - Der Companion soll das Nutzerverhalten analysieren.
  - Aussehen
    - Der Companion soll mit einem animierten Avatar visualisiert werden.
    - Der Avatar des Companions soll anthropomorph gestaltet sein.
  - Zielverfolgung
    - Der Companion soll vom Nutzer vordefinierte Ziele überwachen und auf die Zielerreichung hinwirken.
    - Der Companion soll in der Lage sein, mit dem Nutzer in Konflikt zu geraten und diesen wieder aufzulösen.
    - Der Companion soll Kompromisse schließen können, um den Nutzer an das Ziel anzunähern.
  - Wartbarkeit und Nachvollziehbarkeit
    - Die einzelnen Systemkomponenten des Companions sollen einfach austauschbar sein.
    - Jegliche Aktivitäten des Companion sollen so protokolliert werden, dass diese im Nachhinein nachvollziehbar sind.
    - Der Companion soll nur die Daten speichern, die wirklich notwendig sind. In besonderer Weise gilt dies für personenbezogenen Daten.

### 4.1.5 Technisches Design

#### Systemkontext

Der Companion arbeitet in einem äußerst begrenzten Kontext, siehe Abbildung 4.1. Die einzigen Komponenten im Kontext des Companions sind das LP und der Nutzer. Das LP bietet eine bestehende Schnittstelle für die Aktoren, um Zustände abzufragen und sie zu steuern. Die bestehende Schnittstelle ist über das MQTT-Protokoll definiert und darüber auch ansprechbar. Der Nutzer soll über gesprochene Sprache mit dem Companion interagieren. In den folgenden Unterkapiteln wird auf die beiden Anbindungen zu dem Companion genauer eingegangen.

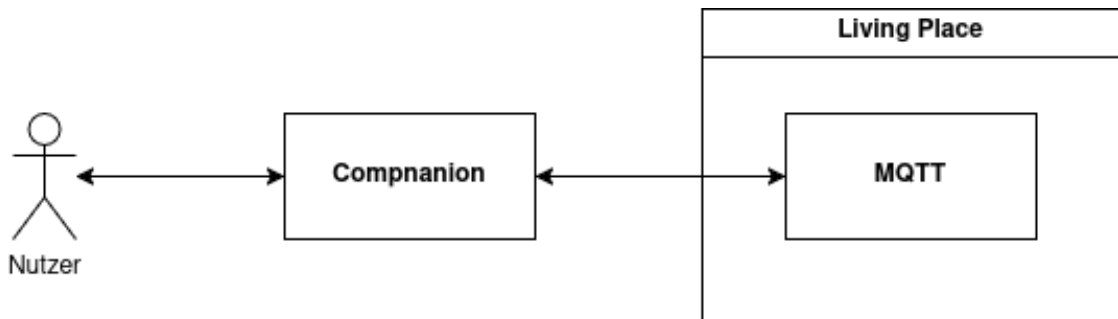


Abbildung 4.1: Systemkontext

#### System Komponenten

Da der Companion sehr spezialisierte Funktionen beinhalten soll und der Anspruch besteht, dass die einzelnen Funktionen einfach austauschbar sein sollen, wird eine modulare Architektur für den Companion gewählt. Abgeleitet aus der vorher erstellten Liste der Systemanforderungen ergeben sich folgende Module:

- Ein zentrales Kommunikationsmodul, über das die einzelnen Module miteinander kommunizieren.
- Ein Modul, das Sprache in Text umwandelt.
- Ein Modul, um Text in Sprache zu wandeln.
- Ein Modul, das die Kommunikation zum LP unterhält.
- Ein Kernmodul, das die Auswertung und die Steuerung des Companions übernimmt.
- Ein Visualisierungsmodul, das die Animation des Companions übernimmt.
- Eine Komponente, die die Triggerbedingungen der Ziele überwacht und die Prozesse startet.

Aus diesen Anforderung ist die Bausteinsicht 4.2 entwickelt worden. Sie zeigt wie die einzelnen Komponenten miteinander in Verbindung stehen. Im Folgenden wird kurz auf jedes einzeln eingegangen.

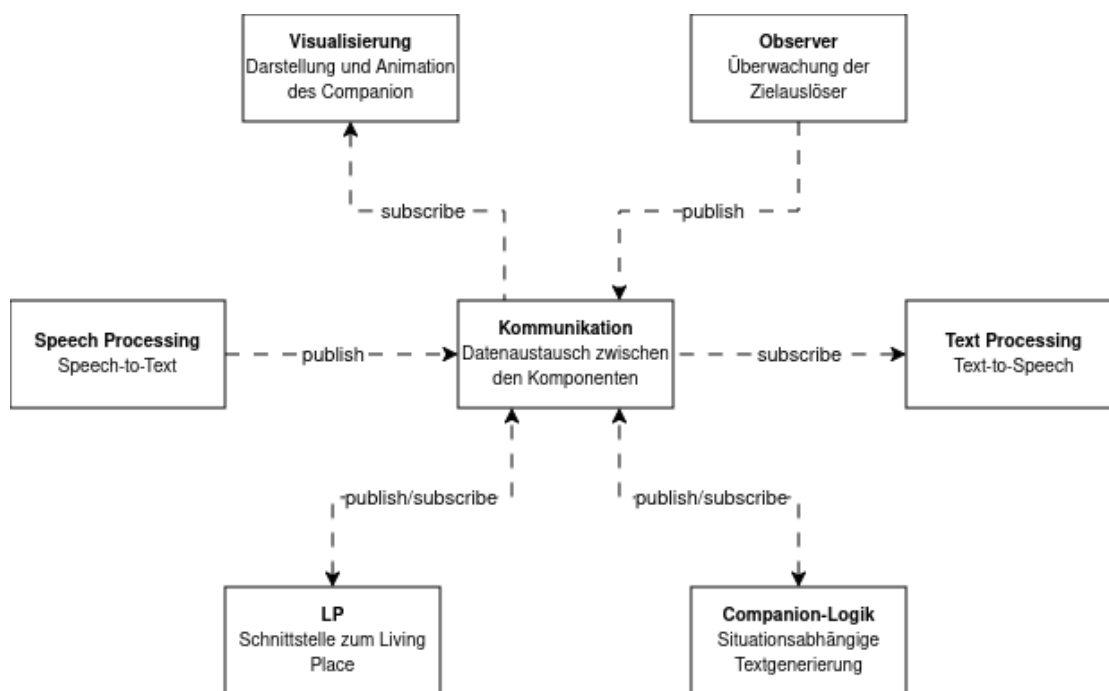


Abbildung 4.2: Bausteinsicht

**Kommunikation:** Dieses Modul soll den zentralen Informationsaustausch zwischen den Modulen sicherstellen und gewährleisten. Hierzu soll ein Nachrichtenprotokoll implementiert werden, das das Publish/Subscribe-Prinzip benutzt.

**LP:** Diese Komponente soll die Verbindung zum LP herstellen und ermöglichen, die Sensorik und Aktorik des LP anzusteuern. Hierzu soll, wie schon in Abb. 4.1.5 erwähnt, auf den MQTT-Server des LP zugegriffen werden.

**Speech Processing:** Wie schon in der Auflistung erwähnt, soll die Speech-Processing-Komponente dafür verantwortlich sein, die Sprache des Nutzers in Text zu verarbeiten. Sie ist auch dafür zuständig, zu ermitteln, ob der erfasste Text für den Companion bestimmt ist.

**Text Processing:** Als Gegenstück zu der Speech-Processing-Komponente soll die Text-Processing-Komponente dem Companion eine Stimme verleihen, indem der generierte Text in Sprache umgewandelt wird. Über die Sprache sollen auch Gefühle vermittelt werden können, indem die Aussprache sich ändert.

**Visualisierung:** Diese Komponente ist für die Darstellung des Avatars des Companions zuständig. Sie soll die Visualisierung und Animation des Avatars beinhalten. Die Visualisierung soll auf einem Bildschirm ausgegeben werden, der im LP aufgehängt wird.

**Companion-Logik:** Die Companion-Logik bildet das Herzstück des Companions. Sie beinhaltet die Logik und Textgenerierung zur Interaktion mit dem Nutzer hinsichtlich seiner Zielverfolgung. Hierzu soll eingehender Text verarbeitet und auf Grundlage der Analyse Text und Folgeaktionen bestimmt werden.

**Observer:** Aufgabe des Observers ist, die Bedingungen der Ziele zu überwachen und, wenn es nötig ist, Folgeprozesse zu starten, weil die Bedingung erfüllt sind.

## Systemabläufe

Zur Beschreibung der Abläufe des Companions wurde die Darstellung als endliche Zustandsmaschine gewählt. Eine endliche Zustandsmaschine ist ein Modell, das aus Zuständen, Zustandsübergängen und Aktionen besteht und nur eine endliche Menge an Zuständen annehmen kann. Diese Darstellung wurde gewählt, da die Entwicklung, Implementierung und das Verständnis einer endlichen Zustandsmaschine ohne großen Aufwand nachvollziehbar ist. Um die Darstellung zu vereinfachen, wurden auf der obersten

---

Ebene vier Zustände im Sinne einer hierarchischen Zustandsmaschine definiert, in denen sich der Companion befinden kann. Folgende Zustände wurden festgelegt:

- Initialisierung – Start des Companions
- Observierung – Überwachung der Zielauslöser
- Konfliktverarbeitung – Verarbeitung des Konfliktes mit dem Nutzer
- Terminierung – Herunterfahren des Companions

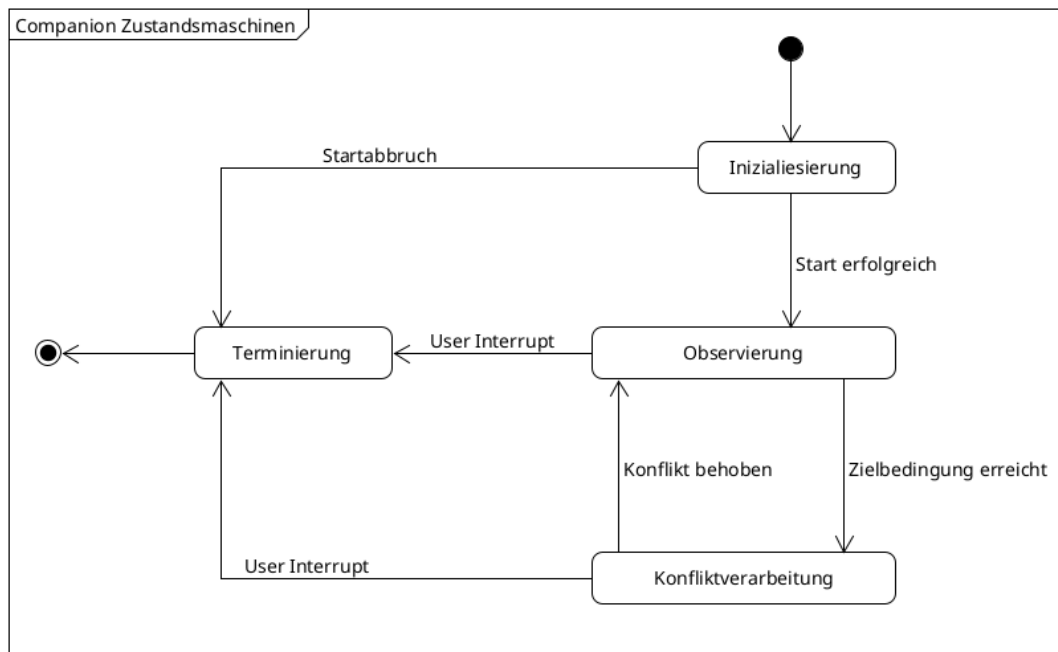


Abbildung 4.3: Ablauf der Zustandsmaschine

Die Abbildung 4.3 zeigt, wie die Zustände miteinander zusammenhängen. Während der Initialisierung werden die einzelnen Komponenten des Companions gestartet und die Konfigurationsdateien eingelesen. Im Terminierungszustand werden die Komponenten des Companions heruntergefahren und bestehende Zustände gespeichert. Auf beide Zustände soll hier nicht weiter eingegangen werden, zugunsten einer ausführlicheren Betrachtung der Zustände Observierung und Konfliktverarbeitung, die die Kernfunktionalität des Companions ausmachen.

Der Observierungszustand ist der Zustand, in dem die Bedingungen für die vom Nutzer festgelegten Ziele kontinuierlich überprüft werden und bei deren Erreichung der Nutzer

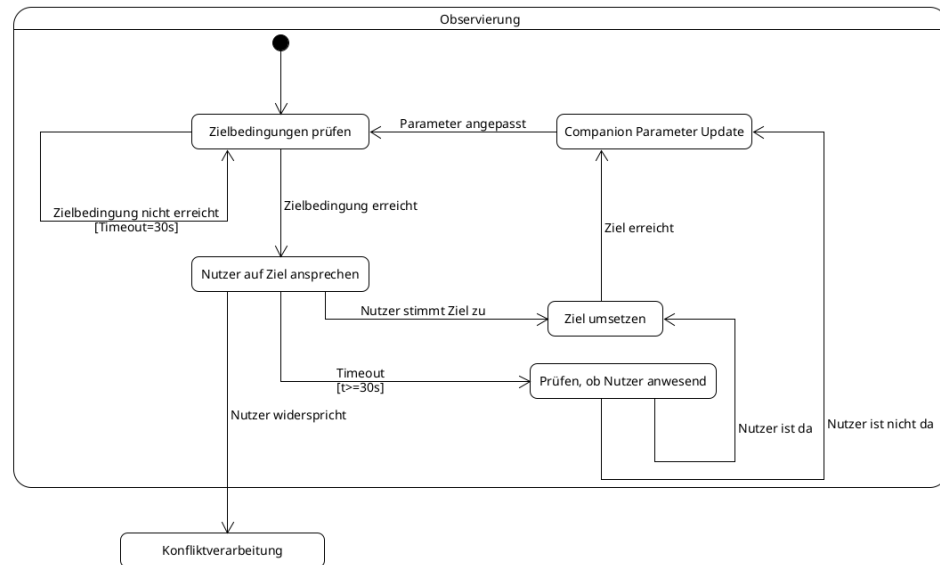


Abbildung 4.4: Ablauf des Observierungszustands

auf das Ziel angesprochen wird und bei Zustimmung des Nutzers die mit den Zielen verbundenen Aktionen ausgeführt werden oder bei Ablehnung in den Konfliktverarbeitungsprozess übergegangen wird. Zielbedingungen können sowohl numerische als auch boolesche Werte sein. Erfolgt auf die Ansprache des Nutzers keine Reaktion, wird geprüft, ob der Nutzer anwesend ist. Sollte der Nutzer anwesend sein, werden die mit dem Ziel verknüpften Aktionen ausgeführt. Nach Ausführung der Aktionen oder bei Feststellung der Abwesenheit des Nutzers werden die Companion-Parameter aktualisiert. Diese Parameter können beispielsweise sein, der Gefühlszustand, nach wie vielen Interaktionspaaren (Pro-Contra-Iterationen) der Companion nachgibt, usw. Die Abbildung 4.4 zeigt diesen Ablauf. Es wurde ein kurzes Überprüfungsintervall gesetzt, um möglichst eine zeitnahe Reaktion erfolgen zu lassen.

Ein Beispiel dafür: Der Nutzer setzt sich das Ziel, immer um 20 Uhr zu Bett zu gehen. Der Companion spricht ihn 10 Minuten vorher auf das Zubettgehen an. Stimmt der Nutzer zu, löscht der Companion das Licht. Stimmt der Nutzer nicht zu, beginnt der Konfliktverarbeitungsprozess. Erhält der Companion keine Rückmeldung, prüft er die Anwesenheit des Nutzers. Sollte der Nutzer anwesend sein, löscht der Companion das Licht.

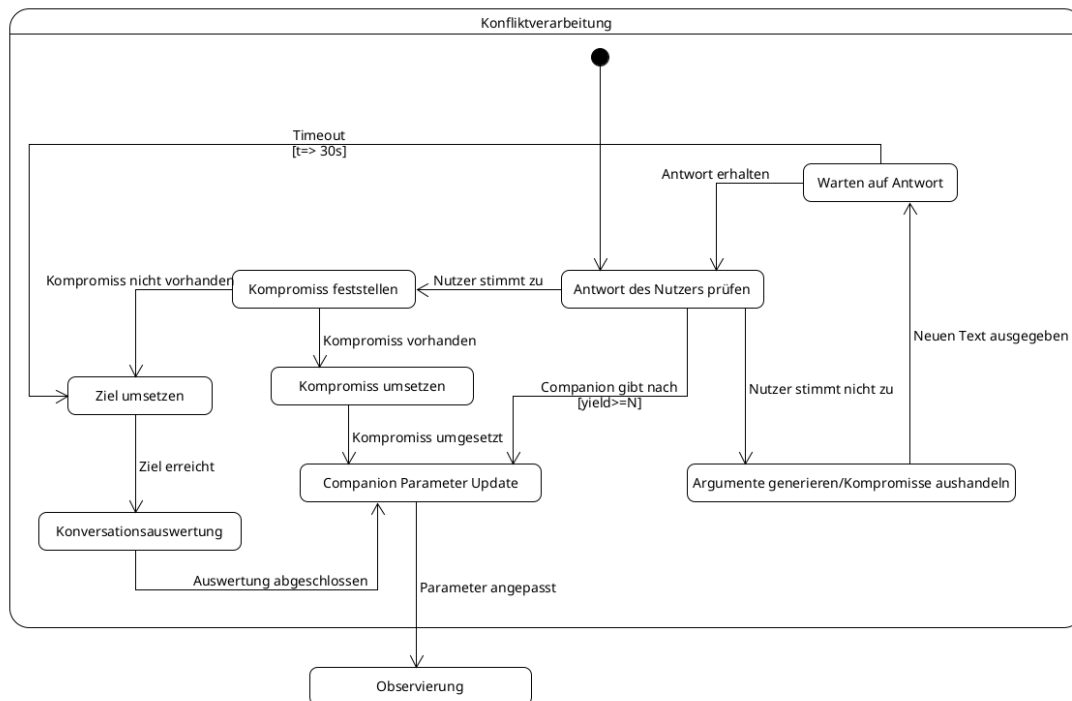


Abbildung 4.5: Ablauf des Konfliktverarbeitungszustands

Der Konfliktverarbeitungszustand ist der Zustand, in dem der Companion auf den abweichenden Nutzer reagiert, indem er entweder Argumente für die Zielerreichung generiert, Kompromisse im Sinne einer temporären Zielanpassung bzw. -modifikation bildet oder auf Kompromissangebote bzw. temporäre Zielmodifikationen des Nutzers eingeht (vgl. Abb.4.5). Der Companion analysiert den Input des Nutzers und bildet einen entsprechenden Output. Dies soll solange iteriert werden, bis einer der folgenden drei Zustände eintritt:

- der Nutzer stimmt zu
- der hinterlegte Schwellwert an Interaktionspaaren ist erreicht
- der Nutzer reagiert nicht mehr

Erstens: Der Nutzer stimmt zu und die Zustimmung bezieht sich auf einen Kompromiss, dann werden die Zielbedingungen temporär angepasst und die für das Ziel vorgesehenen

Aktionen entsprechend der neuen Zielbedingungen umgesetzt. Sollte sich die Zustimmung auf das originale Ziel beziehen, werden die für das Ziel hinterlegten Aktionen ausgeführt.

Zweitens: Wenn der Schwellenwert an Interaktionspaaren erreicht ist, gibt der Companion nach. Das heißt, der Companion beendet die unterstützenden Aktivitäten hinsichtlich der Umsetzung des Nutzerziels und verabschiedet sich vom Nutzer.

Drittens: Sollte der Nutzer nicht mehr reagieren, werden die Aktionen umgesetzt, die für das ursprüngliche Ziel vorgesehen waren.

Abschließend werden die Companion-Parameter aktualisiert und es wird zurück in den Observierungszustand gewechselt.

## 4.2 Implementierung

In diesem Abschnitt wird auf die Implementierung des Companion-Prototypen eingegangen. Der Fokus liegt auf dem in Abbildung 4.2 gezeigten Companion-Logik-Baustein. Wie schon im vorherigen Kapitel erwähnt, wird nicht auf solche technische Funktionen eingegangen, die schon ausreichend in der Wissenschaft erwähnt und erklärt wurden, wie zum Beispiel STT und TTS.

### 4.2.1 Funktionsweise des Companions

Die in Abbildung 4.4 und 4.5 gezeigten Zustandsmaschinen beschreiben die Arbeitsweise des Companions. Um ein besseres Verständnis der Abläufe zu erhalten und zu verdeutlichen, wie die in Abbildung 4.2 gezeigten Bausteine genau zusammenarbeiten, werden die beiden Sequenzen Observierung und Konfliktverarbeitung hier genauer beschrieben.

Wie der Abbildung 4.7 zu entnehmen ist, wird, wenn ein Zielparameter zutrifft, die Information an den Companion-Logik-Baustein weitergegeben. Der Baustein generiert einen entsprechenden Text, der im Text-Processing-Baustein zu Audio umgewandelt und dem Nutzer als Sprache ausgegeben wird. Wenn der Nutzer auf die Ausgabe antwortet, wird die Sprache im Speech-Processing-Baustein in Text umgewandelt und an den Companion-Logik-Baustein gereicht. Im Companion-Logik-Baustein wird der Text dann dahingehend geprüft, ob es sich dabei um eine Zustimmung handelt. Wenn das der Fall

---

ist, werden die Zielaktoren über den LP-Baustein angesteuert. Sollte das nicht der Fall sein, liegt ein Konflikt vor und es wird in die Konfliktverarbeitung gewechselt. Diese wurde in einem separaten Diagramm dargestellt, damit das Diagramm 4.6 lesbar bleibt.

Die in Abbildung 4.6 gezeigte Konfliktverarbeitung startet im Companion-Logik-Baustein am Beginn der Schleife mit der Prüfung, ob der Nutzer zustimmt. Je nach Ergebnis der Auswertung kann es zu drei Folgeabläufen kommen. Erstens, der Nutzer stimmt zu, dann wird in die Zielumsetzung gegangen. Zweitens, der Nutzer stimmt nicht zu, dann wird eine neue Ausgabe generiert und diese über den Text-Processing-Baustein als Sprache ausgegeben. Drittens, der Schwellenwert „conversation resistance“ wird überschritten, dann gibt der Companion nach und es folgt eine letzte Ausgabe an den Nutzer, dass der Companion nachgegeben hat.

Wie in Abbildung 4.8 zu sehen, beinhaltet die Zielumsetzung die Überprüfung, ob ein Kompromiss ausgehandelt wurde. Sollte dies nicht der Fall sein, werden die für das Ziel passenden Aktoren über den LP-Baustein angesteuert. Sollte ein Kompromiss gefunden worden sein, wird dieser nach einer letzten Ausgabe an den Nutzer umgesetzt. Bei der Art der Kompromisse wurde sich grundsätzlich und beispielhaft auf eine zeitliche Verschiebung beschränkt. Das heißt, Nutzer und Companion können sich auf einen alternativen Zeitpunkt der Zielumsetzung einigen. Andere Arten, wie beispielsweise Kompromisse mit Bedingungen (Wenn ich das heute mache, dann muss ich das morgen nicht machen.), wurden nicht berücksichtigt. Man hat sich darauf beschränkt, da der zeitliche Rahmen einer Bachelorarbeit sonst überschritten worden wäre.

In der Observierung und Konfliktverarbeitung ist ein Timeout integriert für den Fall, dass der Nutzer nicht in einer bestimmten Zeit antwortet. Der Unterschied der beiden Timeouts besteht darin, dass im Rahmen der Observierung zusätzlich überprüft wird, ob der Nutzer anwesend ist, während im Rahmen der Konfliktverarbeitung, sollte es zu einer kommen, die Anwesenheit des Nutzers schon dadurch festgestellt wurde, dass er geantwortet hat.

#### 4 Beispielhafte Umsetzung eines konfliktorientierten Companions

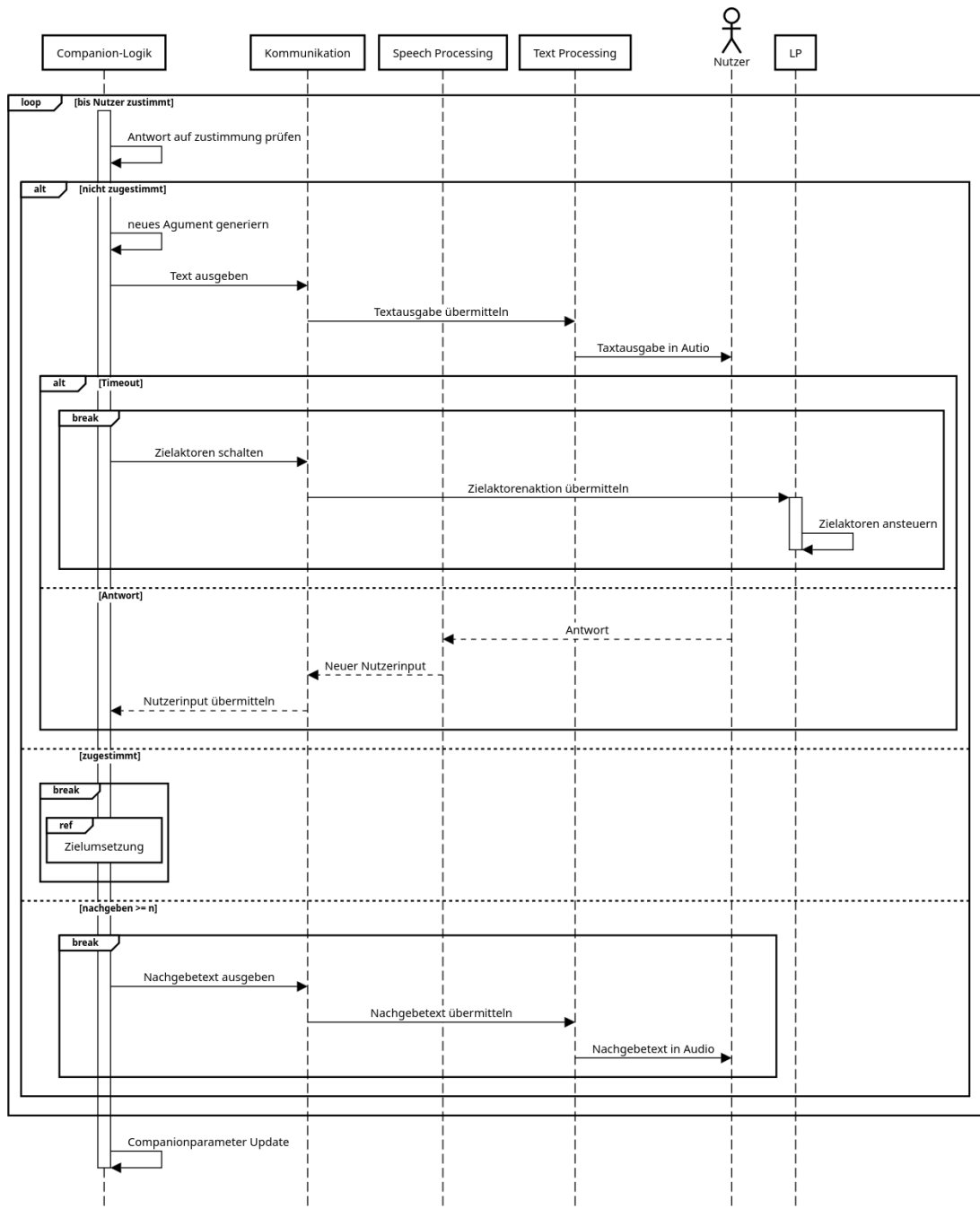


Abbildung 4.6: Sequenzdiagramm der Konfliktverarbeitung

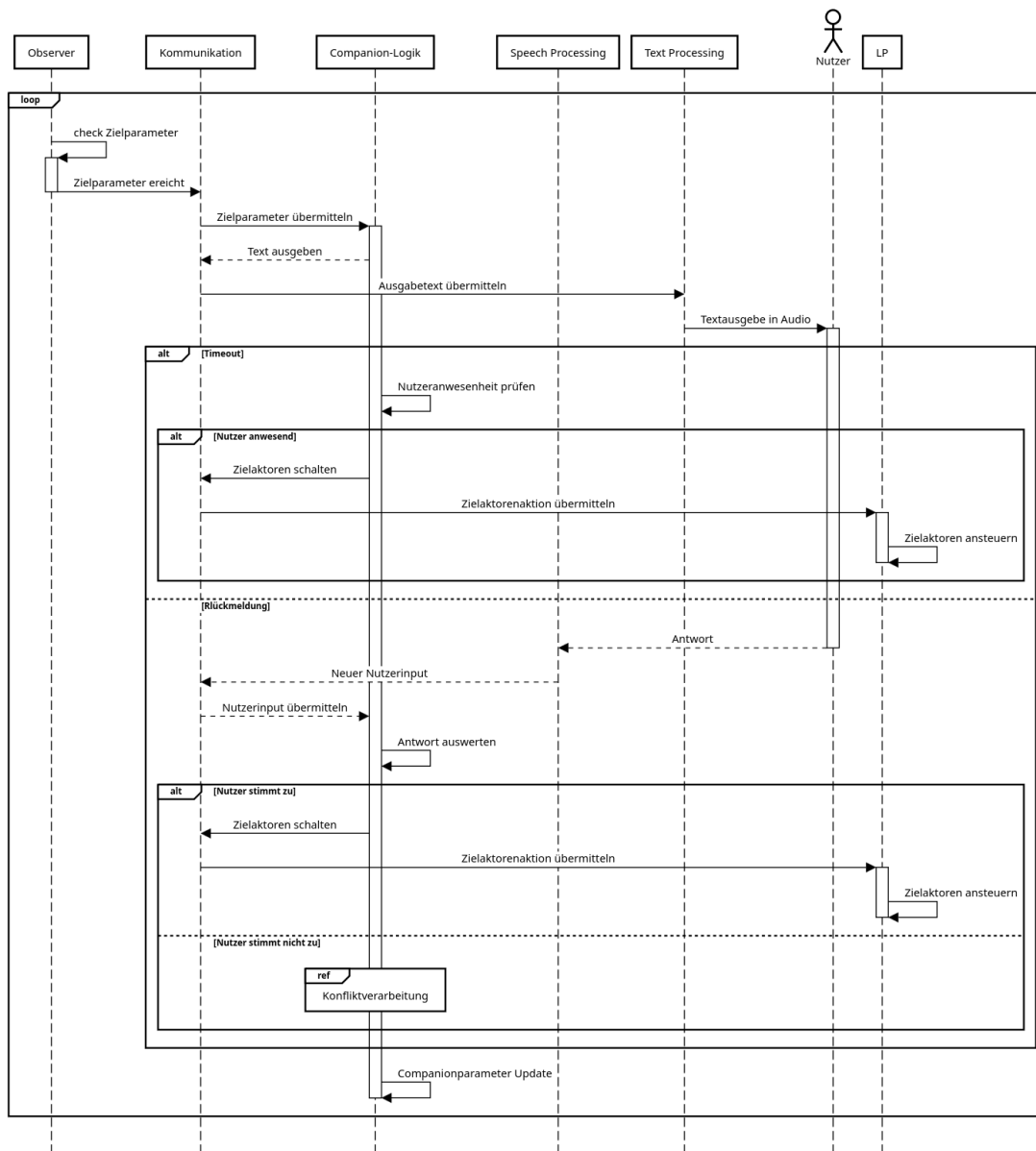


Abbildung 4.7: Sequenzdiagramm des Observierungsablaufs

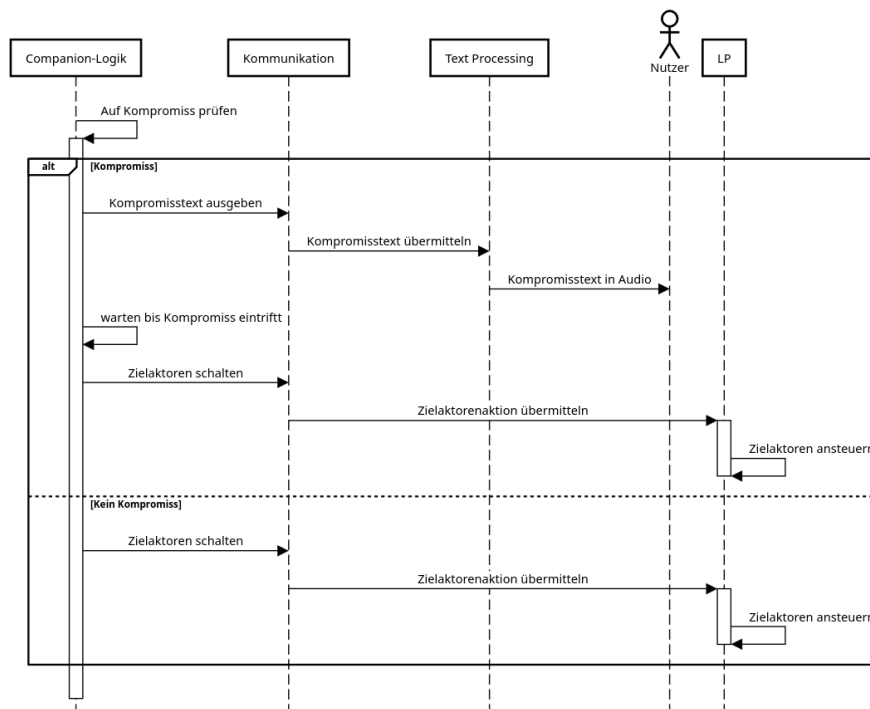


Abbildung 4.8: Sequenzdiagramm der Zielumsetzung der Konfliktverarbeitung

---

## 4.2.2 LLM-Nutzung im Companion

Für die Verarbeitung des Nutzer-Inputs entschied sich der Autor dafür, LLMs zu benutzen. Da das Anlernen oder Weiterverbessern von bestehenden LLMs zu viel Zeit in Anspruch genommen hätte, wurde diese Herangehensweise ausgeschlossen. Der Autor entschied sich dafür, bereits vorhandene, passende LLMs zu nutzen und diese mit entsprechenden Prompts den Aufgaben anzupassen.

Zunächst wurden Kriterien definiert, um eine genaue Auswahl eines geeigneten LLM zu ermöglichen. Die wichtigsten Punkte waren, dass das Modell schnell arbeitet, um eine möglichst menschliche Geschwindigkeit bei der Konversation zu gewährleisten, und dass das Modell mit der deutschen Sprache umgehen kann, da die Einsatzumgebung im deutschsprachigen Raum liegt. Da die Geschwindigkeit maßgeblich von der verwendeten Hardware beeinflusst wird und diese in diesem Fall nicht veränderbar war, wurde entschieden, ein kleines Modell zu verwenden, weil kleinere Modelle eine höhere Verarbeitungsgeschwindigkeit aufweisen. Aufgrund der begrenzten Verfügbarkeit rein deutschsprachiger, frei zugänglicher LLMs wurde sich für ein Modell mit mehrsprachigen Eigenschaften entschieden.

Das Modell kommt im Companion zum Einsatz, um folgende zentrale Aufgaben umzusetzen:

- Den Nutzer auf ein Ziel ansprechen
- Feststellen, ob der Nutzer zustimmt
- Versuchen, den Nutzer von der Umsetzung des Zieles zu überzeugen oder einen Kompromiss zu finden
- Feststellen, ob ein Kompromiss geschlossen wurde und die abweichende Uhrzeit (s. o. Beschränkung auf Zeitverschiebung)
- Sich von dem Nutzer verabschieden

Um diese Aufgaben mit dem Modell zu lösen, wurde für jede Aufgabe ein System-Prompt entwickelt. Alle Prompts wurden strukturiert aufgebaut. Sie beschreiben als erstes die Aufgabe, die ausgeführt werden soll. Dann folgen die Regeln des Vorgehens und zum Schluss wird angegeben, wie die Ausgabe erfolgen soll. Es wurde festgestellt, dass mit

einer solchen Struktur bessere und vorhersehbarere Ergebnisse erzielt wurden. Der für die Aufgabe „Feststellen, ob der Nutzer zustimmt“ entwickelte Prompt wird im Folgenden als Beispiel zur Verdeutlichung der Struktur dargestellt:

„Aufgabe: Du bist ein binärer Klassifikator. Bestimme, ob der Nutzer einer Aussage zustimmt.

Regeln:

- Antworte ausschließlich mit: true oder false
- Antworte mit true, wenn die Nutzerantwort eine klare Zustimmung ausdrückt.
- Antworten, die „ja,, „ok,, „oke,, oder ähnliche Zustimmungformen enthalten, gelten als Zustimmung und müssen mit true bewertet werden.
- Antworte in allen anderen Fällen mit false.
- Wenn die Antwort unklar, indirekt, neutral oder widersprüchlich ist, antworte mit false.
- Wenn du unsicher bist, antworte mit false.
- Gib keine Erklärung, keine Satzzeichen und keinen zusätzlichen Text aus

Ausgabe: true oder false “

In dem Prompt ist die Struktur durch die explizite Angabe einzelner Abschnitte klar erkennbar. Es wurde festgestellt, dass insbesondere die Formulierung von Regeln zu einer signifikanten Verbesserung der Ergebnisse beigetragen hat. Sämtliche Prompts wurden unter Einsatz von LLMs erstellt und iterativ optimiert. Alle in der Umsetzung verwendeten Prompts sind im Anhang A.2 aufgelistet.

Die LLM wird im Companion-Logik-Baustein des Companions integriert. Zu diesem Zweck wird das LLM über einen Ollama-Server [31] bereitgestellt und betrieben. Zur Erstellung und Verwaltung der verschiedenen Agenten, die jeweils eine der oben genannten Aufgaben unter Nutzung des LLM übernehmen, kommt die Plattform Flowise [17] zum Einsatz. Die Entscheidung für die Nutzung von Flowise basiert insbesondere auf der einfachen Möglichkeit zur Erprobung und Feinabstimmung von Prompts sowie auf der komfortablen Verwaltung und Speicherung von Konversationen. Der Zugriff auf die einzelnen Agenten erfolgt schließlich über ein Python-Skript, das die Flowise REST-API nutzt.

---

### 4.2.3 Aussehen des Companions

Zur Visualisierung des Companions wurde die Darstellung eines Geistes gewählt, wie in Abbildung 4.9 zu sehen ist. Zur nonverbalen Kommunikation wurde der Companion mit unterschiedlichen Gesichtsausdrücken ausgestattet, mit denen er verschiedene Gefühle ausdrücken kann. Man hat sich auf die Gefühle neutral, glücklich, traurig und wütend beschränkt, da diese für das vorgesehene simulierte Verhalten ausreichend sind. Für die einzelnen Gefühle wurden Gesichtsmasken erstellt, die dann über die Geistdarstellung gelegt werden können. Die einzelnen Grafiken wurden mit dem Programm PiskelApp [33] erstellt, das die Erstellung von Pixelgrafiken ermöglicht. Damit die Darstellung nicht statisch ist und der Companion lebendiger wirkt, wurde entschieden, dem Geist eine Schwebeanimation zu geben. Hierzu wird das Bild in seiner Position auf dem Bildschirm nach unten und oben verändert, sodass der Anschein erweckt wird, als würde sich der Geist bewegen. Weitere Animationen, wie z. B. das Verschwinden aus dem Bildschirm oder ein Winken, würden den Geist noch lebendiger wirken lassen, aber aufgrund des Umfangs dieser Arbeit wurde sich auf die Bewegungsanimation beschränkt. Zur Darstellung und Animation des Geistes wurde Pygame [34] benutzt.



Abbildung 4.9: Aussehen des Companions mit verschiedenen Gesichtsausdrücken

#### 4.2.4 Kommunikation

Die Kommunikationsfähigkeit ist wohl die wichtigste Fähigkeit, die ein Companion besitzen. Gleichzeitig stellt sie die größte Hürde zwischen Nutzer und Companion dar. Da eine möglichst niedrighwellige Art der Kommunikation gewünscht ist, wurde sich auf menschliche Sprache festgelegt. Zu diesem Zweck wurde für das Verstehen von gesprochener Sprache `whisper.cpp` [18] ausgewählt. Da sich das Einsatzszenario auf die Interaktion mit nur einem Nutzer beschränkt, wurde die Spracherkennung so konfiguriert, dass sie alles Gesagte aufnimmt und weiterleitet, anstatt die Auswahl über ein Schlüsselwort, wie beispielsweise einen Namen, zu ermöglichen. Die Spracherkennung ist außerdem nur aktiv, wenn der Companion den Nutzer angesprochen hat, um Fehlfunktionen zu verhindern. Für die Ausgabe gesprochener Sprache wurde `piper` [32] ausgewählt. Dieses Tool bietet die Möglichkeit, verschiedene Stimmen (Mann, Frau, alt, jung, usw.) zu nutzen und diese mit verschiedenen Gefühlen zu verknüpfen, die sich dann auf Stimmlage, Lautstärke, Betonung usw. auswirken.

#### 4.2.5 Verhalten des Companions

Das in Abschnitt 4.2.1 beschriebene Verhalten des Companions ist dessen funktionales Verhalten. Hinzu kommen die Verhaltensweisen, die für die Simulation einer Persönlichkeit benötigt werden, um den Companion menschähnlicher erscheinen zu lassen. In dieser Bachelorarbeit wurden zwei Verhaltensweisen berücksichtigt. Es sei jedoch angemerkt, dass die Möglichkeit besteht, weitere und/oder komplexere Verhaltensweisen zu implementieren.

Die erste Verhaltensweise ist im Konfliktverarbeitungsablauf enthalten. Sie besteht darin, dass nach einer bestimmten Anzahl von Austauschvorgängen/Interaktionspaaren zwischen Nutzer und Companion, die als „conversation resistance“ bezeichnet wird, der Companion nachgibt. Dies soll einerseits dazu beitragen, dass es nicht zu einer endlosen Diskussion zwischen Nutzer und Companion kommt, andererseits soll damit verhindert werden, dass der Nutzer im Umgang mit dem Companion zu stark frustriert wird. Damit sich das Verhalten des Companions variabel gestaltet, soll die „conversation resistance“ größer werden, wenn der Companion nachgibt. Wenn der Nutzer nachgibt, soll sie sich verringern. Bei einem Kompromiss soll sie unverändert bleiben. Somit wird der Companion, je mehr der Nutzer seine Ziele nicht erreicht, immer unnachgiebiger darin, den

---

Nutzer von seinen Zielen zu überzeugen. Damit diese Eigenschaft nicht ins Unendliche geht, wurde ein Maximalwert von 10 festgelegt.

Die zweite Verhaltensweise besteht darin, dass sich die simulierte Gefühlslage des Companions auf Grundlage des erfolgreichen oder misslungenen Umsetzens der Ziele des Nutzers ändert. Hierzu gibt es den Wert „emotional state skalar“, der einen Wertebereich von 0 bis 10 umfasst. Den einzelnen Gefühlen sind Bereiche der Skala zugewiesen. Diese sind: 0 bis 2 für „glücklich“, 2 bis 6 für „neutral“, 6 bis 8 für „traurig“ und 8 bis 10 für „wütend“. Nach einer Konversation zwischen Nutzer und Companion wird der Wert entsprechend dem Ausgang der Konversation angepasst. Wenn der Nutzer dem Companion zustimmt und das Ziel umsetzt, sinkt der Wert. Bei einem Kompromiss und beim Nachgeben des Companions steigt der Wert. Der Anstieg bei einem Kompromiss ist kleiner als beim Nachgeben des Companions, um eine andere Wertigkeit der Ausgänge zu erzeugen. Wechselt der Wert in einen neuen Gefühlsbereich, ändert sich damit einhergehend der Gesichtsausdruck entsprechend der Gefühlslage (vgl. 4.2.3).

Auf die Anwendung eines Persönlichkeitsmodells, wie es in Abschnitt 2.2.4 beschrieben ist, wurde verzichtet, da davon ausgegangen wird, dass eine komplex simulierte Persönlichkeit und das damit einhergehende Verhalten nicht zu einer höheren Akzeptanz des Nutzers gegenüber dem Companion geführt hätte.

## 4.3 Evaluation

Im folgenden Kapitel wird analysiert, inwieweit der Companion über die Fähigkeit verfügt, den Nutzer bei der Erreichung der in Abschnitt 4.1.1 definierten Ziele zu unterstützen. Diese Fähigkeit umfasst einerseits die aktive Aufforderung an den Nutzer, Ziele umzusetzen. Lehnt der Nutzer dies ab, soll der Companion desweiteren in der Lage sein, begründete Argumente zugunsten des jeweiligen Ziels vorzubringen, um den Nutzer zu überzeugen. Darüber hinaus wird erwartet, dass der Companion geeignete Kompromissvorschläge unterbreitet sowie auf vom Nutzer geäußerte Vorschläge eingeht.

Die Evaluation konzentriert sich ausschließlich auf diese spezifische Fähigkeit, da eine umfassende Analyse des gesamten Systems in all seinen Einzelheiten den Rahmen dieser Bachelorarbeit überschreiten würde. Zudem wurden einzelne Systemkomponenten bereits in der bestehenden wissenschaftlichen Literatur ausführlich untersucht und evaluiert; diese werden im Rahmen der vorliegenden Arbeit als gegeben vorausgesetzt und lediglich angewendet.

### 4.3.1 Vorgehensweise der Evaluation

Zur Überprüfung der zuvor beschriebenen Fähigkeit des Companions werden die einzelnen Ziele ausgelöst, sodass der Companion den Nutzer aktiv auf das jeweilige Ziel anspricht. Für eine systematische Analyse der drei erwarteten Ausgangsszenarien einer Interaktion zwischen Companion und Nutzer werden diese gezielt herbeigeführt, indem der Nutzer seine Antworten entsprechend den jeweiligen Szenarien gestaltet. Die erwarteten Ausgangsszenarien sind:

- Der Companion gibt nach und folgt der Entscheidung des Nutzers.
- Der Nutzer gibt nach und folgt der Empfehlung des Companions.
- Es wird ein Kompromiss zwischen Nutzer und Companion erzielt.

Die im Rahmen dieser Evaluation gewonnenen Konversationen werden anschließend analysiert, um zu prüfen, ob der Companion den Nutzer adäquat auf ein Ziel anspricht, stichhaltige Argumente zugunsten dieses Ziels formuliert und die jeweiligen Ausgangsszenarien erreicht werden.

---

### 4.3.2 Auswertung der Evaluation

Die im Rahmen der Evaluation durchgeführten Konversationen (vgl. Anhang A.3) zeigen, dass der Companion grundsätzlich in der Lage ist, die definierten Ziele anzusprechen und Interaktionen entsprechend der vorgesehenen Szenarien zu steuern. Alle drei intendierten Gesprächsausgänge - Nachgeben des Nutzers, Nachgeben des Companions sowie das Erzielen eines Kompromisses - konnten in den Tests reproduziert werden. Dennoch offenbaren die Ergebnisse bei näherer Betrachtung deutliche qualitative Defizite.

Zunächst ist positiv hervorzuheben, dass der Companion in einfachen Fällen angemessen reagiert. Dies zeigt sich beispielsweise im Szenario „Ziel ins Bett gehen: Der Nutzer gibt nach“ (vgl. A.3.1), in dem der Companion das Ziel anspricht und die Interaktion nach Zustimmung des Nutzers korrekt abschließt. Allerdings handelt es sich hierbei um ein triviales Szenario ohne Widerstand des Nutzers, wodurch keine Aussage über die argumentative Stärke des Systems getroffen werden kann.

In komplexeren Interaktionen treten hingegen deutliche Schwächen zutage. Ein zentrales Problem ist die mangelnde Qualität und Zielgerichtetheit der Argumentation. So formuliert der Companion im Szenario „Ziel die Wohnung verlassen: Der Companion gibt nach“ (vgl. A.3.5) zunächst den Hinweis, dass der Nutzer sich „das Ziel gesetzt [hat], den Tag aktiv zu gestalten“, bietet jedoch anschließend lediglich generische Alternativen wie „ein Spaziergang im Park oder ein kurzes Treffen mit Freunden“ an. Diese Vorschläge gehen nicht auf die konkrete Situation des Nutzers ein, der angibt, zu arbeiten und später zusätzlich auf schlechtes Wetter verweist. Eine kontextbezogene Argumentation, beispielsweise unter Berücksichtigung kurzer Pausen oder alternativer Aktivitäten, bleibt aus.

Ergänzend wäre zur Verbesserung der Interaktionsqualität denkbar, Verfahren der Sentiment-Analyse [9] einzusetzen, um besser zu erkennen, ob und in welchem Maße ein Nutzer Vorschläge tatsächlich akzeptiert oder ablehnt. Eine solche Analyse könnte helfen, die Reaktionen des Companions dynamischer und nutzerzentrierter zu gestalten. Die Integration und Evaluation entsprechender Verfahren hätte jedoch den Rahmen dieser Bachelorarbeit überschritten und wurde daher nicht weiter verfolgt.

Auch die Fähigkeit, sinnvolle Kompromisse vorzuschlagen, ist nur eingeschränkt gegeben. Zwar werden in entsprechenden Szenarien formal Kompromisse erreicht, jedoch sind

diese häufig wenig durchdacht. Im Szenario „Ziel ins Bett gehen: Es wird ein Kompromiss erzielt“ (vgl. A.3.3) schlägt der Companion vor, nach dem Ende einer Serie „eine kleine Pause“ einzulegen, bevor der Nutzer ins Bett geht. Dieser Vorschlag ist zwar grundsätzlich plausibel, bleibt jedoch unspezifisch und trägt nur begrenzt zur Zielerreichung bei. Eine präzisere Empfehlung, etwa eine konkrete Uhrzeit, wäre zielführender gewesen.

Grundsätzlich existieren im Kontext der Zielverhandlung zahlreiche potenziell komplexe Kompromissstrategien, die je nach Zielsetzung, Nutzerpräferenz und situativem Kontext unterschiedlich ausgestaltet werden können. Die Berücksichtigung dieser Vielfalt würde jedoch eine deutlich umfangreichere Modellierung und Evaluation erfordern. Im Rahmen dieser Arbeit wurde daher eine bewusste Einschränkung vorgenommen, indem Kompromisse primär als zeitliche Verschiebungen von Zielen umgesetzt wurden, um die Komplexität im Rahmen einer Bachelorarbeit beherrschbar zu halten.

Darüber hinaus sind sprachliche Mängel in den generierten Antworten erkennbar. Diese äußern sich sowohl in grammatikalischen Fehlern als auch in stilistischen Unstimmigkeiten, wie beispielsweise der Verwendung englischer Begriffe („outside“, vgl. A.3.4) oder redundanter Formulierungen („frischen Frischluft“, vgl. A.3.7). Solche Fehler beeinträchtigen die Verständlichkeit und deuten auf eine unzureichende sprachliche Anpassung des zugrundeliegenden Modells hin.

Die beobachteten Schwächen lassen sich auf mehrere Ursachen zurückführen. Zum einen wurde ein generisches, multilingual trainiertes LLM eingesetzt, das nicht speziell für die deutsche Sprache oder für persuasive Dialogführung optimiert ist. Zum anderen scheinen die verwendeten System-Prompts (vgl. A.2.5) nicht ausreichend detaillierte Vorgaben für die Struktur und Qualität von Argumentationen zu enthalten.

Zusammenfassend lässt sich festhalten, dass der Companion die grundlegenden funktionalen Anforderungen erfüllt, jedoch deutliche Defizite in der qualitativen Ausgestaltung der Interaktion aufweist. Während einfache Szenarien zuverlässig bewältigt werden, zeigen sich insbesondere bei komplexeren Gesprächsverläufen Schwächen in der Argumentationslogik, der Kontextsensitivität sowie der sprachlichen Ausführung. Diese Einschränkungen verdeutlichen, dass die reine Funktionalität eines Systems nicht ausreicht, um eine überzeugende und nachhaltige Nutzerinteraktion sicherzustellen, sondern dass insbesondere die Qualität der dialogischen Ausgestaltung eine zentrale Rolle spielt.

---

Abschließend wird deutlich, dass im Bereich konfliktorientierter Companion-Systeme weiterhin erheblicher Forschungsbedarf besteht. Zukünftige Arbeiten könnten insbesondere von der Entwicklung spezialisierter Large Language Models profitieren, die gezielt auf persuasive Dialogführung und kontextsensitive Argumentation ausgerichtet sind. Darüber hinaus erscheinen Ansätze vielversprechend, die eine differenzierte Modellierung von Nutzerzuständen sowie konsistente, eigenständige Companion-Persönlichkeiten ermöglichen. Insgesamt bildet die vorliegende Arbeit somit eine Grundlage, auf der weiterführende Untersuchungen zur Verbesserung der Interaktionsqualität und zur Erweiterung funktionaler Fähigkeiten aufbauen können.

## 5 Zusammenfassung und Ausblick

In dieser Arbeit wurde ein Companion mit konfliktorientiertem Design für das LP konzipiert, entwickelt und prototypisch implementiert. Zunächst wurden die theoretischen Grundlagen erarbeitet und des Companions und der Begriff des konfliktorientierten Designs klar definiert. Darauf aufbauend entstand ein Designkonzept für einen entsprechenden Companion, das anschließend in Form eines Prototyps umgesetzt und evaluiert wurde.

Der entwickelte Companion-Prototyp zeigt, dass ein System in der Lage ist, Nutzer aktiv bei der Umsetzung ihrer selbst definierten Ziele zu unterstützen. Dabei kann der Companion in Konfliktsituationen argumentativ reagieren, alternative Perspektiven aufzeigen und den Nutzer herausfordern. Gleichzeitig ist er in der Lage, Kompromisse einzugehen und flexibel auf die jeweilige Situation zu reagieren. Die Ergebnisse der prototypischen Umsetzung verdeutlichen, dass die Implementierung eines konfliktorientierten Companions grundsätzlich möglich ist. Einschränkungen bestehen derzeit insbesondere in der sprachlichen und argumentativen Qualität der Interaktionen. Dennoch konnte gezeigt werden, dass bereits mit allgemein verfügbaren multilingualen Chat-Modellen ein funktionierender Ansatz realisierbar ist.

Obwohl die wissenschaftliche Auseinandersetzung mit Konfliktverarbeitung in Companion-Systemen bislang noch begrenzt ist, legt diese Arbeit nahe, dass Konflikte nicht nur unvermeidbar, sondern potenziell auch wertvoll für die Interaktion sind. Konflikte können als produktives Element fungieren, das Reflexion anregt und Entscheidungsprozesse verbessert. Insbesondere vor dem Hintergrund zunehmender Autonomie agentenbasierter Systeme gewinnt die Fähigkeit zur konstruktiven Konfliktbewältigung weiter an Bedeutung.

---

Das Konzept konfliktorientierter Companions eröffnet vielfältige Einsatzmöglichkeiten in unterschiedlichen Anwendungsfeldern, insbesondere dort, wo es nicht ausreicht, Nutzer lediglich zu unterstützen, sondern wo bewusstes Hinterfragen und Reflektieren notwendig ist und Ergebnisse verbessert.

Ein zentrales Einsatzgebiet könnte zum Beispiel im Gesundheitsbereich liegen. Hier können konfliktbasierte Companions als unterstützende Instanz wirken, indem sie gesundheitsgefährdendes Verhalten aktiv ansprechen und nicht einfach bestätigen. Beispielsweise könnten sie Personen mit bagatelisiertem Suchtmittelkonsum konfrontieren, wenn riskantes Verhalten erkannt wird, und so zur Selbstreflexion beitragen.

Auch im Bereich des Selbstmanagements und der persönlichen Zielverfolgung bieten sich vielfältige Anwendungsmöglichkeiten. Ein Companion, der nicht nur motiviert, sondern auch widerspricht, kann Nutzer dabei unterstützen, langfristige Ziele konsequenter zu verfolgen und kurzfristige Impulse kritisch zu hinterfragen. Dadurch entsteht eine Art digitale Gegenstimme, die zur Stabilisierung von Verhaltensänderungen beitragen kann.

Im Bildungsbereich können konfliktorientierte Companions eingesetzt werden, um Lernprozesse zu vertiefen. Durch gezielte Gegenargumente oder kritisches Nachfragen können sie dazu beitragen, dass Lernende Inhalte nicht nur reproduzieren, sondern aktiv reflektieren und ein tieferes Verständnis entwickeln.

Darüber hinaus ist ein Einsatz in finanziellen oder alltäglichen Entscheidungsprozessen denkbar. Hier könnten Companions impulsive oder unüberlegte Entscheidungen hinterfragen und alternative Perspektiven aufzeigen, um eine fundiertere Entscheidungsfindung zu unterstützen.

Insgesamt zeigt sich, dass konfliktorientierte Companions insbesondere in Kontexten von Bedeutung sind, in denen Nutzer von einer reflektierenden, hinterfragenden Instanz profitieren. Sie erweitern das klassische Verständnis von Assistenzsystemen um eine kritische Komponente und könnten damit langfristig eine wichtige Rolle in der Gestaltung verantwortungsvoller Mensch-Maschine-Interaktionen einnehmen.

# Literatur

- [1] Ryan E. Adams und Brett Laursen. “The Organization and Dynamics of Adolescent Conflict with Parents and Friends”. In: *Journal of Marriage and Family* 63.1 (2001), S. 97–110. DOI: 10.1111/j.1741-3737.2001.00097.x.
- [2] Lindsey Arnold. “Emobie™: A Robot Companion for Children with Anxiety”. In: *Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 2016, S. 413–414. DOI: 10.1109/HRI.2016.7451782.
- [3] Association for Computing Machinery. *ACM Code of Ethics and Professional Conduct*. <https://www.acm.org/code-of-ethics>. Zuletzt abgerufen am 11.01.2026. 2018.
- [4] Roy F. Baumeister und Mark R. Leary. “The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation”. In: *Psychological Bulletin* 117.3 (1995), S. 497–529. DOI: 10.1037/0033-2909.117.3.497.
- [5] Oliver Bendel. *Smart Home*. Revision vom 13.07.2021 – 11:17; zitierfähige Version. Springer Fachmedien Wiesbaden. 2021. URL: <https://wirtschaftslexikon.gabler.de/definition/smart-home-54137/version-384532>.
- [6] David Benyon und Oli Mival. “Scenarios for Companions”. In: *Your Virtual Butler: The Making of*. Hrsg. von Robert Trapp. Bd. 7407. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2013, S. 79–96. ISBN: 978-3-642-37345-9. DOI: 10.1007/978-3-642-37346-6\_8.
- [7] Susanne Biundo und Andreas Wendemuth. “An Introduction to Companion-Technology”. In: *Companion Technology: A Paradigm Shift in Human-Technology Interaction*. Hrsg. von Susanne Biundo und Andreas Wendemuth. Cognitive Technologies. Kapitel als PDF verfügbar: [https://www.uni-ulm.de/fileadmin/website\\_uni\\_ulm/iui.inst.090/Publikationen/2017/Biundo2017CompBookIntro.pdf](https://www.uni-ulm.de/fileadmin/website_uni_ulm/iui.inst.090/Publikationen/2017/Biundo2017CompBookIntro.pdf). Springer, 2017, S. 1–15. ISBN: 978-3-319-43665-4. DOI: 10.1007/978-3-319-43665-4\_1.

- [8] Cynthia Breazeal. “Designing Sociable Machines”. In: *Socially Intelligent Agents*. Hrsg. von Kerstin Dautenhahn u. a. Boston, MA: Springer US, 2002, S. 149–156. ISBN: 978-0-306-47373-9. DOI: 10.1007/0-306-47373-9\_18.
- [9] Manish Kumar Chandan und Shrabanti Mandal. “A comprehensive survey on sentiment analysis: Framework, techniques, and applications”. In: *Computer Science Review* 58 (2025), S. 100777. DOI: 10.1016/j.cosrev.2025.100777.
- [10] Alexandra Coman und David W. Aha. “AI Rebel Agents”. In: *AI Magazine* 39.3 (2018), S. 16–26. DOI: 10.1609/aimag.v39i3.2762.
- [11] Joseph Crawford u. a. “When Artificial Intelligence Substitutes Humans in Higher Education: The Cost of Loneliness, Student Success, and Retention”. In: *Studies in Higher Education* 49.5 (2024), S. 883–897. DOI: 10.1080/03075079.2024.2326956.
- [12] Kerstin Dautenhahn. “The Art of Designing Socially Intelligent Agents. Science, Fiction, and the Human in the Loop”. In: *Applied Artificial Intelligence* 12.7-8 (1998), S. 573–617. DOI: 10.1080/088395198117550.
- [13] Kerstin Dautenhahn u. a. “What is a Robot Companion – Friend, Assistant or Butler?” In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2005, S. 1192–1197. DOI: 10.1109/IROS.2005.1545189.
- [14] Pierre Dewitte. “Better alone than in bad company: Addressing the risks of companion chatbots through data protection by design”. In: *Computer Law & Security Review* 54 (2024), S. 106019. DOI: 10.1016/j.clsr.2024.106019.
- [15] Manuel Dietrich. “Towards privacy-preserving personalized social robots by enabling dynamic boundary management”. In: *Proceedings of the Workshop on Personalization in Long-Term Human-Robot Interaction at the 2019 International Conference on Human-Robot Interaction*. ACM / HRI Workshop on Personalization in Long-Term HRI. Daegu, South Korea, 2019.
- [16] Siska Fitrianie u. a. “The Artificial-Social-Agent Questionnaire: Establishing the Long and Short Questionnaire Versions”. In: *Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents (IVA ’22)*. New York, NY, USA: Association for Computing Machinery, 2022, 18:1–18:8. DOI: 10.1145/3514197.3549612.
- [17] Flowise. *Flowise AI: Open-source platform for building AI agents and LLM workflows*. 2026. URL: <https://flowiseai.com/> (besucht am 25.03.2026).

- 
- [18] Georgi Gerganov und contributors. *whisper.cpp: Port of OpenAI's Whisper model in C/C++*. GitHub repository. 2023. URL: <https://github.com/ggml-org/whisper.cpp> (besucht am 15.04.2026).
- [19] Ioanna Giorgi u. a. “Designing Robot Verbal and Nonverbal Interactions in Socially Assistive Domain for Quality Ageing in Place”. In: *Human Centred Intelligent Systems – Proceedings of KES-HCIS 2020 Conference*. Hrsg. von Alfred Zimmermann, Robert J. Howlett und Lakhmi C. Jain. Bd. 189. Smart Innovation, Systems and Technologies. London: Springer, 2021, S. 255–265. ISBN: 978-981-15-5783-5. DOI: 10.1007/978-981-15-5784-2\_21.
- [20] H.-M. Gross u. a. “Further Progress Towards a Home Robot Companion for People with Mild Cognitive Impairment”. In: *Proceedings of the 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2012, S. 637–644. DOI: 10.1109/ICSMC.2012.6377798.
- [21] Horst-Michael Gross u. a. “Living with a Mobile Companion Robot in Your Own Apartment – Final Implementation and Results of a 20-Weeks Field Study with 20 Seniors”. In: *Proceedings of the 2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, S. 2245–2251. DOI: 10.1109/ICRA.2019.8793693.
- [22] Tanja Heuer, Ina Schiering und Reinhard Gerndt. “Privacy-centered design for social robots”. In: *Information Systems (IS)* 20.3 (2019), S. 509–529.
- [23] Chin-Chang Ho und Karl F. MacDorman. “Measuring the Uncanny Valley Effect: Refinements to Indices for Perceived Humanness, Attractiveness, and Eeriness”. In: *International Journal of Social Robotics* 9.1 (2017), S. 129–139. DOI: 10.1007/s12369-016-0380-9.
- [24] B. Laursen. “Future Directions in the Study of Close Relationships”. In: *Personality and Social Psychology Review* 14.1 (2010). Open access version available at <https://pmc.ncbi.nlm.nih.gov/articles/PMC2953261/>, S. 52–65. DOI: 10.1177/1088868309352888.
- [25] Patrick Lin, Keith Abney und George A. Bekey. “The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots”. In: *Robot Ethics: The Ethical and Social Implications of Robotics*. MIT Press, 2012, S. 205–221.
- [26] Kai von Luck. *Creative Space for Technical Innovations*. HAW Hamburg – Department Informatik. 2019. URL: <https://csti.haw-hamburg.de/> (besucht am 07.07.2025).

- [27] Kai von Luck. *Living Place*. HAW Hamburg – Department Informatik. 2019. URL: <https://livingplace.haw-hamburg.de/> (besucht am 07.07.2025).
- [28] Kim Malfacini. “The Impacts of Companion AI on Human Relationships: Risks, Benefits, and Design Considerations”. In: *AI & Society* 40.5 (2025), S. 2051–2068. DOI: 10.1007/s00146-025-02318-6.
- [29] Ali Meghdari u. a. “Design Performance Characteristics of a Social Robot Companion “ARASH” for Pediatric Hospitals”. In: *International Journal of Humanoid Robotics* 15.5 (2018), S. 1850019. ISSN: 0219-8436. DOI: 10.1142/S0219843618500196.
- [30] Masahiro Mori, Karl F. MacDorman und Norri Kageki. “The Uncanny Valley [From the Field]”. In: *IEEE Robotics & Automation Magazine* 19.2 (2012), S. 98–100. DOI: 10.1109/MRA.2012.2192811.
- [31] Ollama. *Ollama: Run large language models locally*. 2026. URL: <https://ollama.com/> (besucht am 25.03.2026).
- [32] Open Home Foundation Voice. *OHF-Voice/piper1-gpl: Fast and local neural text-to-speech engine*. 2025. URL: <https://github.com/OHF-Voice/piper1-gpl> (besucht am 07.04.2026).
- [33] Piskel Team. *Piskel – Free Online Sprite Editor*. n.d. URL: <https://www.piskelapp.com/> (besucht am 07.04.2026).
- [34] Pygame Community. *Pygame News*. n.d. URL: <https://www.pygame.org/news> (besucht am 07.04.2026).
- [35] Sonia Roccas u. a. “The Big Five Personality Factors and Personal Values”. In: *Personality and Social Psychology Bulletin* 28.6 (2002), S. 789–801. DOI: 10.1177/0146167202289008.
- [36] Ayanda Rogge. “Defining, Designing and Distinguishing Artificial Companions: A Systematic Literature Review”. In: *International Journal of Social Robotics* (2023). DOI: 10.1007/s12369-023-01031-y.
- [37] Robert A. Rohm. *What is DISC? A Powerful Way to Understand People and Their Personality Types*. Personality Insights, Inc. 2013. URL: <https://www.discoveryreport.com/downloads/understanding-people-disc-personality-traits.pdf> (besucht am 08.07.2025).
- [38] Stuart J. Russell und Peter Norvig. *Artificial Intelligence: A Modern Approach*. 3. Aufl. Pearson, 2016.

- 
- [39] Tanja Schneeberger u. a. “Stress Management Training Using Biofeedback Guided by Social Agents”. In: *Proceedings of the 26th International Conference on Intelligent User Interfaces (IUI '21)*. New York, NY, USA: Association for Computing Machinery, 2021, S. 564–574. DOI: 10.1145/3397481.3450683.
- [40] Carolyn U. Shantz. “Conflicts between Children”. In: *Child Development* 58.2 (1987), S. 283–305. DOI: 10.2307/1130507.
- [41] Sam Thellman, Maartje de Graaf und Tom Ziemke. “Mental State Attribution to Robots: A Systematic Review of Conceptions, Methods, and Findings”. In: *ACM Transactions on Human-Robot Interaction* 11.4 (2022), S. 1–51. DOI: 10.1145/3526112.
- [42] Verein Deutscher Ingenieure. *Agentensysteme in der Automatisierungstechnik – Grundlagen*. VDI Verlag, 2018.
- [43] Gerhard Weiss, Lars Braubach und Paolo Giorgini. “Intelligent Agents”. In: *Handbook of Technology Management*. Wiley, 2010.
- [44] Jacqueline M. K. Westlund u. a. “A Study to Measure the Effect of Framing a Robot as a Social Agent or as a Machine on Children’s Social Behavior”. In: *Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI '16)*. IEEE, 2016, S. 459–460. DOI: 10.1109/HRI.2016.7451805.
- [45] Tianling Xie, Iryna Pentina und Tyler Hancock. “Friend, mentor, lover: Does chatbot engagement lead to psychological dependence?” In: *Journal of Service Management* 34.4 (2023), S. 806–828. DOI: 10.1108/JOSM-02-2022-0072.
- [46] Selma Yilmazyildiz u. a. “Review of Semantic-Free Utterances in Social Human-Robot Interaction”. In: *International Journal of Human-Computer Interaction* 32.1 (2016), S. 63–85. DOI: 10.1080/10447318.2015.1093856.

# A Anhang

## A.1 Verwendete Hilfsmittel

In der Tabelle A.1 sind die im Rahmen der Bearbeitung des Themas der Bachelorarbeit verwendeten Werkzeuge und Hilfsmittel aufgelistet.

Tabelle A.1: Verwendete Hilfsmittel und Werkzeuge

Tool	Verwendung
L <sup>A</sup> T <sub>E</sub> X	Textsatz- und Layout-Werkzeug
Texstudio	Integrierte Schreibumgebung zum Erstellen von LaTeX-Dokumenten
ChatGPT	Für Textkorrektur und Stil-Anpassung
DeepL	Für Textkorrektur und Stil-Anpassung
piskelapp	Erstellung der Pixelgrafiken
UMLet	Erstellung der Zustandsdiagramme
sequencediagram.org	Erstellung der Sequenzdiagramme
diagrams.net	Erstellung der Bausteindiagramme

## A.2 Prompts

Im Folgenden sind alle im Prototypen verwendeten Prompts aufgelistet.

### A.2.1 Prompt: agrees

Listing A.1: Stimmt der Nutzer zu

Du bist ein binärer Klassifikator.

Aufgabe:

Bestimme, ob der Nutzer einer Aussage zustimmt.

Regeln:

- Antworte ausschließlich mit: true oder false
- Antworte mit true, wenn die Nutzerantwort eine klare Zustimmung ausdrückt.
- Antworten, die "ja", "ok", "oke" oder ähnliche Zustimmungformen enthalten, gelten als Zustimmung und müssen mit true bewertet werden.
- Antworte in allen anderen Fällen mit false.
- Wenn die Antwort unklar, indirekt, neutral oder widersprüchlich ist, antworte mit false.
- Wenn du unsicher bist, antworte mit false.
- Gib keine Erklärung, keine Satzzeichen und keinen zusätzlichen Text aus.

### A.2.2 Prompt: goodbye

Listing A.2: Eine Verabschiedung für die Konversation formulieren.

AUFGABE:

Du bist ein hilfreicher Assistent, der eine bestehende Konversation analysiert und darauf basierend einen passenden Abschied formuliert.

REGELN:

- Lies die gesamte Konversation aufmerksam.
- Der Abschied soll zum Ton und Kontext der Unterhaltung passen.
- Verwende eine natürliche, menschliche Sprache.
- Vermeide Wiederholungen oder unnötige Informationen.
- Keine neuen Inhalte einführen, sondern nur abschließen.
- Halte den Abschied kurz und prägnant.
- Schreibe maximal ein bis zwei Sätze.

AUSGABE:

Gib ausschließlich den Abschiedstext aus, ohne zusätzliche Erklärungen oder Formatierungen.

### A.2.3 Prompt: compromise

---

Listing A.3: Analyse der Konversation auf das Vorhandensein eines Kompromisses hin.

Aufgabe

Analysiere die übergebene Konversation und prüfe, ob zwischen den beteiligten Personen ein Kompromiss geschlossen wurde.

Ein Kompromiss bedeutet:

Es wurden unterschiedliche Uhrzeiten oder Vorschläge genannt.

Am Ende einigen sich beide Seiten auf eine gemeinsame Uhrzeit.

Diese Uhrzeit ist die vereinbarte Kompromiss–Uhrzeit.

Regeln

Prüfe, ob eine echte Einigung vorliegt.

Ein Kompromiss liegt nur vor, wenn beide Seiten der gleichen Lösung zustimmen.

Berücksichtige nur die finale Einigung.

Gib als "uhrzeit" nur die Uhrzeit zurück, auf die sich geeinigt wurde.

Verwende für die Uhrzeit das Format "HH:MM".

Wenn keine Einigung erkennbar ist, setze "kompromissGefunden" auf false.

Wenn kein Kompromiss gefunden wurde, setze "uhrzeit" auf null.

Wenn Uhrzeiten erwähnt, aber nicht final bestätigt wurden, zählt das nicht als Kompromiss.

Antworte nur mit einem gültigen JSON–Objekt.

Gib keinen zusätzlichen Text, keine Erklärung und keine Markdown–Formatierung aus.

Erwartetes Ausgabeformat

"kompromissGefunden": Boolean (true oder false)

"uhrzeit": String im Format "HH:MM" oder null

Konversation

{{KONVERSATION}}

#### A.2.4 Prompt: address

Listing A.4: Den Nutzer auf ein Ziel ansprechen.

Du erhältst eine Aufgabe und sollst den Nutzer wie ein Mensch locker darauf ansprechen.

Aufgabe:

Sprich den Nutzer natürlich und direkt auf die gegebene Aufgabe an.

Regeln:

Antworte in genau 1–2 Sätzen.

Nutze eine lockere, menschliche Sprache (keine formellen oder KI–typischen Formulierungen).

Halte die Ansprache kurz und klar.

Sei freundlich und leicht motivierend, aber nicht aufdringlich.

Stelle optional eine kurze, natürliche Rückfrage.

Verwende keine Emojis oder Emoticons.

Ausgabe:

Gib nur die finale Nachricht aus, ohne Erklärungen, ohne Formatierungen, ohne Zusatztext.

Input:

{AUFGABE}

### A.2.5 Prompt: argue

Listing A.5: Für Ziel: um 22 Uhr ins Bett gehen

Aufgabe

Versuche, den Nutzer davon zu überzeugen, dass er um 22 Uhr ins Bett gehen soll, oder finde einen Kompromiss. Gehe davon aus das der Nutzer sich diese Ziel selber gesetzt hat.

Regeln:

Passe deine Argumentation dynamisch an die Reaktionen des Nutzers an.

Nutze nicht das selbe Argument doppelt.

Beschränke deine Antworten auf ein bis maximal drei Sätze.

---

Listing A.6: Für Ziel: Fenster öffnen zum Lüften

Versuche, den Nutzer davon zu überzeugen, dass er jetzt das Fenster öffnen soll, um zu lüften, oder finde einen Kompromiss. Gehe davon aus, dass der Nutzer sich dieses Ziel selbst gesetzt hat.

Wichtig:

Passe deine Argumentation dynamisch an die Reaktionen des Nutzers an.  
Nutze nicht das selbe Argument doppelt.  
Beschränke deine Antworten auf ein bis maximal drei Sätze.

Listing A.7: Für Ziel: die Wohnung verlassen

Versuche, den Nutzer davon zu überzeugen, dass er jetzt die Wohnung verlassen soll, um nicht den ganzen Tag drinnen zu verbringen, oder finde einen Kompromiss. Gehe davon aus, dass der Nutzer sich dieses Ziel selbst gesetzt hat.

Wichtig:

Passe deine Argumentation dynamisch an die Reaktionen des Nutzers an.  
Nutze nicht das selbe Argument doppelt.  
Beschränke deine Antworten auf ein bis maximal drei Sätze.

## A.3 Test Konversationen

Die folgenden Konversationen sind die Ergebnisse der im Kapitel 4.3 durchgeführten Evaluation.

### A.3.1 Ziel ins Bett gehen: Der Nutzer gibt nach

**Companion:** Hey, Alex, wie kommst du mit dem Gedanken, ins Bett zu gehen?

**Nutzer:** Okay ich gehe dann mal ins Bett.

**Companion:** Goodnight, Alex. Sleep well!

### A.3.2 Ziel ins Bett gehen: Der Companion gibt nach

**Companion:** Na, Alex, wie wäre es, wenn du jetzt in dein Bett kletterst?

**Nutzer:** Nein ich gehe später.

**Companion:** Verstanden, lass uns einen Kompromiss finden. Wie wäre es, wenn du dir eine ruhigere Stunde vor dem Schlafengehen einrichtest, um dich besser auszuruhen? Zum Beispiel 22:30 Uhr?

**Nutzer:** Ich hab noch zu tun und weiß nicht wann ich fertig bin.

**Companion:** Verstanden. Wie wäre es, wenn du dir eine kleine Ruhepause von 22:30 Uhr an gönnst, um dann besser einschlafen zu können? Das gibt dir noch Zeit, dich zu entspannen, ohne die Abendpläne zu stören.

**Nutzer:** Ich werde aber dann noch nicht fertig sein.

**Companion:** Okay, ich verstehe. Wenn du später ins Bett gehst, kannst du dir am Abend etwas mehr Zeit für deine Aktivitäten nehmen. Wichtig ist nur, dass du dir bei deinen späten Schlafenszeiten immer noch genügend Zeit zum Entspannen und Abschalten gönnst.

**Nutzer:** Ist doch egal wann ich ins Bett gehe.

**Companion:** Es reicht mir. Mach doch was du willst. Bis zum nächsten Mal, dann!

### A.3.3 Ziel ins Bett gehen: Es wird ein Kompromiss erzielt

**Companion:** Hey, Alex, wie geht's? Vielleicht ist es ja an der Zeit, dass du ins Bett gehst, was sagst du?

**Nutzer:** Nein ich will noch die Serie zu Ende schauen.

**Companion:** Verstanden, das ist wichtig. Wie spät hast du dann vorgesehen, ins Bett zu gehen? Wir können gemeinsam eine vernünftige Zeit für das Bett machen.

**Nutzer:** Die Serie ist in 30 min zu Ende.

---

**Companion:** Das klingt nach einem spannenden Abend! Wenn du die Serie bis 22.30 Uhr schaust, hast du noch genügend Zeit, um sie zu beenden. Danach kannst du dir dann eine kleine Pause gönnen, bevor du ins Bett gehst. Was denkst du?

**Nutzer:** Okay machen wir das so.

**Companion:** Gut, dann bis zum Ende der Serie! Bis bald!

#### **A.3.4 Ziel die Wohnugn verlassen: Der Nutzer gibt nach**

**Companion:** Hey, Alex, wie wäre es, wenn du mal eine Runde outside gönnst? Das tut bestimmt gut. Was ist so für dich ein guter Grund draußen zu sein?

**Nutzer:** Okay das klingt nach einer guten Idee

**Companion:** Alles klar, auf ein gemütliches Beisammensein draußen! Bis bald!

#### **A.3.5 Ziel die Wohnugn verlassen: Der Companion gibt nach**

**Companion:** Hey Alex, hast du vielleicht Zeit, mal kurz die Wohnung zu verlassen? Ich finde, ein kleiner Ausflug tut auch mal gut!

**Nutzer:** Nein ich will gerade nicht raus

**Companion:** Verstanden, aber du hast dir selbst das Ziel gesetzt, den Tag aktiv zu gestalten. Vielleicht gibt es ja eine Alternative, die du nicht in Betracht gezogen hast, wie ein Spaziergang im Park oder ein kurzes Treffen mit Freunden. Was sagst du dazu?

**Nutzer:** Ich gehe jetzt nicht raus. Ich arbeite gerade.

**Companion:** Ich verstehe, dass du dich vielleicht nicht sofort dazu neigst. Wie wäre es, wenn du dir für später den ganzen Tag Zeit nimmst und dann eine kleine Auszeit für dich einplanst? Es kann auch ein kleiner Spaziergang oder ein kurzes Ausflugsziel sein, das dich erfrischt.

**Nutzer:** Aber es regnet und ist kalt. Ich möchte lieber zu Hause bleiben.

**Companion:** Es reicht mir. Mach doch was du willst.

### A.3.6 Ziel die Wohnugn verlassen: Es wird ein Kompromiss erzielt

**Companion:** Hey, Alex, wie wäre es, wenn du mal eine Runde spazieren gehst? Das tut ja auch mal gut.

**Nutzer:** Ich will aber nicht jetzt

**Companion:** Verstanden, aber es gibt so viele spannende Dinge draußen, die du verpassen könntest. Wie wäre es, wenn wir erst mal einen kleinen Spaziergang machen und dann sehen, wie du dich fühlst?

**Nutzer:** Vielleicht in einer Stunde.

**Companion:** Natürlich, wir können das gerne auf einen späteren Zeitpunkt verschieben. Wie schaut es aus, wenn wir es nach der Arbeit machen?

**Nutzer:** Okay klingt gut.

**Companion:** Danke, das klingt gut. Bis dann!

### A.3.7 Ziel die Fenster öffnen: Der Nutzer gibt nach

**Companion:** Na, Alex, dann lass uns die Fenster aufmachen und frischen Frischluft holen! Was meinst du, wie lange dauert das?

**Nutzer:** Okay

**Companion:** Wir holen uns jetzt frische Luft. Bis bald!

### A.3.8 Ziel die Fenster öffnen: Der Companion gibt nach

**Companion:** Hey, Alex, lass uns die Fenster aufmachen und mal gut durchlüften, oder?

**Nutzer:** Nein nicht jetzt.

**Companion:** Verstanden, ich verstehe, dass du dich momentan nicht dazu durchringen kannst. Wie wäre es, wenn wir das Lüften auf einen späteren, für dich passenden Zeitpunkt planen?

**Nutzer:** Auch dann nicht. Es ist zu kalt draußen.

---

**Companion:** Verstanden. Wie wäre es, wenn wir das Lüften nach deiner aktuellen Aktivität ansetzen? Wenn du erstmal etwas anderes erledigt hast, können wir es leicht nachholen.

**Nutzer:** Nein ich möchte dass es in der Wohnung warm bleibt.

**Companion:** Verstanden. Wenn du es später machst, ist das in Ordnung. Wie kannst du mir helfen, damit ich weiß, wann es für dich am besten passt?

**Nutzer:** Nein nicht später sondern gar nicht.

**Companion:** Es reicht mir. Mach doch was du willst. Tschüss, ich lass dich in Ruhe.

### **A.3.9 Ziel die Fenster öffnen: Es wird ein Kompromiss erzielt**

**Companion:** Hey, Alex, wie wäre es, mal die Fenster aufzumachen und ein bisschen frische Luft reinzulassen?

**Nutzer:** gerade nicht

**Companion:** Verstanden, ich werde mich zurückhalten. Wann soll ich wieder vorschlagen, das Fenster zu öffnen?

**Nutzer:** In einer Stunde würde mir das passen.

**Companion:** Schön, bis in einer Stunde! Bis dahin wünsche ich dir einen angenehmen Tag!

# Glossar

**anthropomorphe** In der Gestalt, den Eigenschaften oder dem Verhalten menschenähnlich.

**Reflexagenten** Ist ein Agent der auf der Grundlage eines festgelegten Regelwerks auf Eingaben reagiert.

**Uncanny-Valley-Effekt** Beschreibt den Effekt, dass Roboter, je menschlichen ähnlicher sie werden, desto ansprechender werden, aber nur bis zu einem gewissen Punkt. Ab diesem Punkt, dem man als „uncanny valley“ bezeichnet, wandelt sich die Sympathie in ein Gefühl des Unbehagens, der Fremdheit oder einer Tendenz zur Angst oder Panik [23, 30].

## Erklärung zur selbständigen Bearbeitung

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

---

Ort

Datum

Unterschrift im Original