SMART COMPANION 2

ETHIKKRITERIEN

für die Entwicklung von technischen Assistenzsystemen für ältere Menschen am Beispiel eines sprechenden Staubsaugerroboters zur Sturzerkennung



Akademie für Altersforschung am
HAUS DER BARMHERZIGKEIT

BOSCH

ARBEITER-SAMARITER-BUND ÖSTERREICHS

University of Applied Sciences St. Pölten

SMART COMPANION
Whitepaper
2025
https://research.ustp.at/projekte/smart-companion-2

Impressum

DESIGN

Stefanie Größbacher

PROJEKTMITARBEITER*INNEN

University of Applied Sciences St. Pölten

David Feilacher
Bernhard Girsule
Stefanie Größbacher
Mario Heller
Arushi Jain
Andreas Jakl
Sabine Lehner
Johannes Pflegerl
Florian Zahorka

Robert Bosch AG

Philip Dimitrov
Sven Dominka
Jakob Hausberger
Henrik Smith

Akademie für Altersforschung am Haus der Barmherzigkeit

Matei Capatu
Thomas Dorner
Christina Fastl

Arbeiter-Samariter-Bund Österreichs, Gruppe Linz

Michael Kaindleinsberger

ETHIKBOARDMITGLIEDER

Daniel Green
Andreas Jakl
Michael Kaindleinsberger
Sabine Lehner
Birgit Meinhard-Schiebel
Angelika Neuer
Johannes Pflegerl
Veronika Schauer
Henrik Smith

ZITIERVORSCHLAG

University of Applied Sciences St. Pölten, Robert Bosch AG, Akademie für Altersforschung am Haus der Barmherzigkeit, Arbeiter-Samariter-Bund Österreichs, Gruppe Linz und Smart-Companion-Ethikboard (2025): Ethikkriterien für die Entwicklung von technischen Assistenzsystemen für ältere Menschen am Beispiel eines sprechenden Staubsaugerroboters zur Sturzerkennung. Projekt „Smart Companion“. Whitepaper.

Akademie für Altersforschung am HAUS DER BARMHERZIGKEIT

BOSCH

ARBEITER-SAMARITER-BUND ÖSTERREICHS

University of Applied Sciences St. Pölten

FFG Forschung wirkt.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie

Abstract

Das vorliegende Whitepaper beleuchtet ethische Fragestellungen in der Entwicklung technischer Assistenzsysteme für ältere Menschen anhand des Projekts **Smart Companion 2**. Im Zentrum steht ein KI-gestützter, sprechender Staubsaugerroboter mit integrierter Sturzerkennung und Sprachassistenz, der für die Verwendung im eigenen Zuhause entwickelt wurde. Dies soll nicht zuletzt dazu beitragen, dass sturzgefährdete bzw. ältere Menschen möglichst lange im eigenen Zuhause leben können. Aufbauend auf einem interdisziplinären Forschungsansatz und unter Einbindung eines projektinternen Ethikboards wurden **sieben zentrale Ethikkriterien** identifiziert und praktisch umgesetzt:

1. ***Sicherheit der Nutzenden,***
2. ***Datenschutz und Privatsphäre,***
3. ***Transparenz und Nachvollziehbarkeit,***
4. ***Achtung der Würde,***
5. ***Autonomie und Nicht-Unterwerfung,***
6. ***Vermeidung von Diskriminierung und Achtung von Diversität***
7. ***Vorbeugung gegen Täuschung und Verwechslung***.

Das Whitepaper verknüpft theoretische Grundlagen mit praktischen Umsetzungserfahrungen und bietet damit eine Orientierung für zukünftige Projekte im Bereich von Assistenzsystemen (für ältere Menschen). Es wird dafür plädiert, ethische Prinzipien und Nutzendenperspektiven bereits im Designprozess zu berücksichtigen, um menschenzentrierte Technikgestaltung und soziale Akzeptanz zu fördern.

Inhaltsverzeichnis

Abstract 3

1 Einleitung 5

2 Zum Projekt Smart Companion 2 7

3 Ethikkriterien: Theoretische Herleitung 11

4 Ethikkriterien: Definition und praktische Umsetzung 17

Kriterium 1: Sicherheit der Nutzenden 19

Kriterium 2: Datenschutz und Privatsphäre 20

Kriterium 3: Transparenz und Nachvollziehbarkeit 22

Kriterium 4: Achtung der Würde der Nutzenden 23

Kriterium 5: Autonomie und Nicht-Unterwerfung 24

Kriterium 6: Vermeidung von Diskriminierung und Achtung von Diversität 24

Kriterium 7: Vorbeugung gegen Täuschung und Verwechslung 26

Literatur 27

1 Einleitung

Ethik in der Technikentwicklung hat in den letzten Jahren angesichts des aktuellen Booms von KI-basierten Anwendungen vermehrt mediale und politische Aufmerksamkeit erhalten. Neben den propagierten Potentialen in verschiedenen Anwendungsbereichen (für private, berufliche bis hin zu bürokratisch-politischen Zwecken) werden auch mögliche Risiken beleuchtet. Supranationale politische Institutionen wie die *UNESCO* oder *EU* haben entsprechende Empfehlungen und Verordnungen verabschiedet (Europäische Union, 2024; UNESCO, 2023). Das, was im Kontext von KI als ethisch problematisch gehandhabt wird – wie die autonome bzw. automatisierte Entscheidung und unvorhersehbare Folgen auf menschliches Leben und Gesellschaften – zeigt allgemein ethische Herausforderungen und Probleme in Bezug auf das Verhältnis zwischen Technologien und Menschen auf. Der Berücksichtigung von ethischen Prinzipien liegen die Annahmen zugrunde, dass Technologien nie neutral sind und ihrerseits bestimmte Werte und gesellschaftliche Realitäten reproduzieren. Folglich können Technologien das Leben von Menschen (potenziell unvorhersehbar tiefgreifend) beeinflussen. Die Orientierung an ethischen Prinzipien bei der Entwicklung oder spätestens bei der Zulassung kann dabei helfen, negative Folgen abzuwenden, wenngleich unerwünschte Konsequenzen in der Techniknutzung nie gänzlich antizipiert werden können. Das Verständnis, was als problematisch oder risikobehaftet gilt und inwiefern Politik und Gesellschaft regulierend in derlei technische Entwicklungen eingreifen soll, die als sog. Innovationsprozesse, technische Fortschritte und nicht zuletzt wirtschaftlich lukrative Entwicklungen verstanden werden, wird jedoch teils strittig verhandelt.

In der Forschung und Entwicklung finden ethikbezogene Überlegungen u.a. in Form von *Ethics by Design* (Brey & Dainow, 2024; Nussbaumer et al., 2021), *Privacy by Design* (Cavoukian, 2011) oder vom *digitalen Humanismus* als wissenschaftlicher Bewegung (Werthner et al., 2019) bereits seit längerem Berücksichtigung. Ethikkriterien in der Technikentwicklung fungieren als normative Leitlinien, um sicherzustellen, dass technologische Systeme u.a. menschenzentriert, verantwortungsvoll und sozial verträglich gestaltet werden. Da Fragen der Ethik nicht nur die Implementierung und den Einsatz im Alltag betreffen, müssen diese bereits im Entwicklungsprozess berücksichtigt werden. Wie im *Ethics-by-Design*-Ansatz vorgeschlagen, betreffen diese durchaus kleinteilige Designentscheidungen.

KAPITEL 1 - EINLEITUNG

Ethikbezogene Fragestellungen werden auch spezifisch für die Entwicklung von technischen Assistenzsystemen für ältere Menschen formuliert, die als Zielgruppe im Zentrum dieses Projektes stehen (Körtner, 2016; Weber, 2015, 2016). Auch im Folgenden geht es um die Entwicklung eines Staubsaugerroboters (**Smart Companion**) zur Sturzerkennung für ältere Menschen. Im zugrundeliegenden Forschungs- und Entwicklungsprojekt **Smart Companion 2** wurden ethikbezogene Aspekte kontinuierlich in der Entwicklung einbezogen.

Das Ziel des vorliegenden Whitepapers ist es, sowohl auf theoretischer Basis, als auch mit Blick auf die Praxis, Anhaltspunkte für Projekte anzubieten, in denen technische Assistenzsysteme für ähnliche Anwendungen und soziale Kontexte entwickelt werden. Daher wird zunächst das Projekt (**Kapitel 2 - Zum Projekt Smart Companion 2**) vorgestellt, anschließend ein Überblick über Ethikkriterien in der Forschungsliteratur gegeben (**Kapitel 3 - Ethikkriterien - Theoretische Herleitung**) und schließlich jene Ethikkriterien präsentiert, die für die Entwicklung technischer Assistenzsystemen für ältere Menschen als besonders relevant identifiziert wurden (**Kapitel 4 - Ethikkriterien - Definition und praktische Umsetzung**). Dabei wird jeweils nach einer allgemeinen Beschreibung die **Umsetzung der Ethikkriterien** am Beispiel von **Smart Companion** vorgestellt.

2 Zum Projekt *Smart Companion 2*

Ziel des FFG-geförderten Forschungs- und Entwicklungsprojekts *Smart Companion 2*[1] ist die Erforschung der sozialen und technischen Anforderungen eines sozial akzeptierten Staubsaugerroboters mit dem Namen ***Smart Companion*** (deutsch: Schlauer Begleiter), der neben der gängigen, den Alltag erleichternden Reinigungsfunktion auch der **Sturzerkennung** sowie **Sturzprävention** dient und eine **Assistenzfunktion** umfasst (siehe Abbildung 1). So ist angedacht, dass das Smart Companion System an bestehende Heimnotruf-Infrastrukturen angebunden wird, um im Bedarfsfall – ähnlich wie bei den bekannten Notrufuhren – die Situation abzuklären und ggf. Hilfe zu schicken.

Das Projekt setzt an dem Problem an, dass bestehende Heimnotrufsysteme für Senior*innen meist eine aktive Kontaktaufnahme erfordern, eine hohe Falsch-Positiv-Rate aufweisen, mit einem Stigma verbunden sind und oft nicht richtig oder selten verwendet werden (Coughlin et al., 2007; Heinbüchner et al., 2010; Huang & Huang, 2020).



***Abbildung 1**: Schematische Darstellung: Anbindung von **Smart Companion** an eine Heimnotrufinfrastruktur und Abklärung des Hilfebedarfs. Nachdem der*die Nutzer*in gestürzt ❶, findet Smart Companion im Zuge seiner Such- oder Reinigungsfahrten die Person und fragt, ob sie Hilfe benötigt ❷. Sofern dies entweder mündlich durch den*die Nutzer*in bestätigt wird oder keine verbale Reaktion durch den*die Nutzer*in erfolgt, meldet Smart Companion diesen Fall an die angebundene Heimnotruforganisation ❸. Mitarbeiter*innen der Heimnotruforganisation klären, ob tatsächlich Hilfe benötigt wird und entsenden im Bedarfsfall ein Rettungsmittel ❹.*

[1] Das Projekt wurde durch das FFG-Förderprogramm „IKT der Zukunft" gefördert. Die Initiative wurde vom Bundesministerium für Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie ins Leben gerufen (2022–2025). In diesem Projekt arbeiteten die University of Applied Sciences St. Pölten (USTP), die Robert Bosch AG (BOSCH), die Akademie für Altersforschung am Haus der Barmherzigkeit (AAF) und der Arbeiter- Samariter-Bund Österreichs, Gruppe Linz (ASB) zusammen.

KAPITEL 2 – ZUM PROJEKT SMART COMPANION 2

Vier Anwendungsfälle bzw. Szenarien werden durch Smart Companion abgedeckt:

(A) Erkennung einer gestürzten Person, die nicht mehr ansprechbar ist, woraufhin die Notruforganisation kontaktiert wird,

(B) Erkennung einer gestürzten Person, die noch ansprechbar ist und Smart Companion mittels verbaler Kommunikation vermitteln kann, ob sie Hilfe benötigt,

(C) aktiver Ruf nach Hilfe durch den*die Nutzer*in, und

(D) Sturzvermeidung: der*die Nutzer*in wird vor potenziellen Hindernissen gewarnt.

Fälle (A), (B) und (D) treten ein, wenn Smart Companion eine von dem*der Nutzer*in eingestellte Reinigungsfahrt oder sogenannte Suchfahrt (zur Überprüfung, ob alles in Ordnung ist) macht. Fall (C) wird proaktiv von dem*der Nutzer*in verbal ausgelöst.

Zentrales Element des Smart Companion (als Prototyp) ist ein handelsüblicher Staubsaugerroboter, welcher mit einer Sprachassistenz (*Amazons Alexa*) als Benutzungsschnittstelle ausgestattet wird, um die Bedienung und Kommunikation zu erleichtern. Mit dem Staubsaugerroboter und der Sprachassistenz werden zwei von Senior*innen als nützlich eingestufte Haushaltsgeräte vereint und für Assistenzzwecke verknüpft. Handelsübliche Staubsaugerroboter fertigen standardmäßig Fotos zur Navigation an. Diese Fotos werden für die Hindernis- und Sturzerkennung mittels zusätzlicher Bildanalyse analysiert. Im Sinne der **Datenminimierung** und zum **Schutze der Privatsphäre** werden sämtliche Fotos bis auf den klar definierten Ausnahmefall, dass eine Person vermutlich gestürzt und nicht mehr ansprechbar ist, automatisch nach der lokal am Roboter erfolgten Bildanalyse gelöscht. Somit werden standardmäßig keine Fotos am Roboter gespeichert oder übermittelt. Sie werden ebenso nicht für weiteres Training des neuronalen Netzes eingesetzt. Im Bedarfsfall (z.B., wenn vermutet wird, dass der*die Bewohner*in gestürzt ist) wird mittels Sprachassistenz geklärt, ob Hilfe bei einer Notruforganisation gerufen werden soll. Sofern die*der Nutzer*in dem vorab zugestimmt hat, wird das Foto, das den Alarm ausgelöst hat, an die Notruforganisation übermittelt, damit sich die Heimnotrufmitarbeiter*innen ein klareres Bild machen können, was die Alarmierung ausgelöst hat.

Über die genannten vier Anwendungsfälle hinweg verfolgte das Projekt die Ziele, den Einrichtungsprozess (d.h., die Konfiguration) bedürfnisorientiert zu gestalten, die Sprachinteraktion möglichst robust zu designen, und die sozialen und technischen Bedingungen zu ermitteln, die eine Gefahrenerkennung mit Unterstützung von Machine Learning Algorithmen unter Berücksichtigung von ethischen Prinzipien erlauben. Um diese Ziele zu verfolgen, kombinierte das Projekt bei der Entwicklung Prinzipien und Methoden des *Ethics-by-Designs* und *Human-Centered-Designs*, um die Nutzer*innen mit ihren Bedürfnissen sowie ethische Standards von Anfang an in der Prototypenentwicklung zu berücksichtigen. Dementsprechend wurden Senior*innen, Angehörige und Professionist*innen aus der Praxis in Gruppendiskussionen, Einzelinterviews, moderierten User*innen-Tests und Pilotfeldstudien (in denen 6 Nutzer*innen jeweils einen Monat lang bei sich zuhause getestet haben) direkt in die Entwicklung einbezogen.

Um zu eruieren, *welche Ethikkriterien* nun auf welche Weise in der Entwicklung zu berücksichtigen sind, wurde einschlägige Forschungsliteratur konsultiert und die ermittelten Kriterien wurden in Abstimmung mit Expert*innen aus der Praxis, Technik und einem projektinternen Ethikboard diskutiert. Das projektinterne Ethikboard setzte sich aus Repräsentant*innen aus den Nutzer*innengruppen, Recht, Technik und einer Heimnotruforganisation zusammen. In den jährlichen Ethikboardsitzungen wurden Antworten auf diverse ethikbezogene Fragen erarbeitet.

So ging es bspw. darum, nach ethischen Kriterien festzulegen, welche Daten für die Anwendungsfälle aufgezeichnet, übermittelt und gespeichert werden, und gleichzeitig Privatsphäre sowie Sicherheit zu gewährleisten. Die im Zuge dieses Prozesses erarbeiteten Kriterien wurden in einem Prototyp umgesetzt und abschließend evaluiert.

In den folgenden beiden Abschnitten werden die ethikbezogenen Ergebnisse des dreieinhalbjährigen Projekts in Form eines kurzen Literaturüberblicks (Kapitel 3) und der Beschreibung der Ethikkriterien auf einer allgemeinen und projektspezifischen Ebene (Kapitel 4) vorgestellt.

KAPITEL 2 - ZUM PROJEKT SMART COMPANION 2

In einem weiteren Schritt wurde im Rahmen des Projektes eine Umfrage mit Menschen im Alter von 65 bis 86 Jahren durchgeführt[2]. In einem repräsentativen Sample von 500 Personen wurde Technikaffinität, Sturzangst, aber auch Akzeptanz oder mögliche Hindernisse der Akzeptanz eines Saugstaubroboters mit Sprachassistenz zur Sturzerkennung erhoben. Eine solche Erhebung ist aus ethischen Gründen besonders sinnvoll, weil sie die Perspektiven und Bedürfnisse älterer Menschen ernst nimmt und sicherstellt, dass technische Innovationen nicht über ihre Köpfe hinweg, sondern im Einklang mit ihren Wünschen und ihrer Lebensrealität entwickelt werden.

[2] Hierfür erfolgte die Einreichung eines Ethikantrags bei der Ethikkommission des Haus der Barmherzigkeit. (Positives Votum vom 21.10.2024. #241021_5.1)

3 Ethikkriterien: Theoretische Herleitung

Wie eingangs erwähnt, liegen Empfehlungen, Verordnungen und wissenschaftliche Publikationen vor, in denen zahlreiche ethische Dimensionen bzw. Prinzipien für die Technikentwicklung – in unterschiedlichen Abstraktionsgraden – formuliert werden. Diese reichen von allgemeinen Werten und Risiken bis hin zu konkreteren Bewertungskriterien. Angesichts der teils abstrakten und überlappenden Werte wie **Menschenwürde**, **Autonomie** und **Gerechtigkeit**, sowie der Einsicht, dass ethisch basierte Entscheidungen stark vom Kontext der jeweiligen Technologie abhängen und daher auf Einzelfallbasis getroffen werden müssen (Manzeschke et al., 2013; Nussbaumer et al., 2021), wird im Folgenden ein knapper Überblick über bestehende ethische Dimensionen gegeben, bevor in Kapitel 4 die im Projekt angewandten Ethikkriterien vorgestellt werden.

Ethikkriterien für die Technikentwicklung basieren sowohl auf **rechtlichen Rahmenbedingungen** als auch gesellschaftlichen **Werten**. Während Rechte, etwa in Form von Gesetzen, eine verbindliche Grundlage für die Entwicklung und Bewertung technischer Systeme bieten, sind ethische Kriterien häufig weniger eindeutig kodifiziert und bedürfen einer Interpretation. Da ethikbasierte Entscheidungen vom jeweiligen Kontext abhängen, können dabei sowohl konfligierende Interessen als auch normative Dilemmata entstehen, die sorgfältige Abwägungsprozesse erfordern (Panico et al., 2020; Weber, 2015). So können sich Spannungsfelder zwischen ethischen Prinzipien wie **Sicherheit** versus **Privatsphäre**[3] ergeben. Neben der kontextabhängigen Abwägung sind auch individuelle Präferenzen zu berücksichtigen, weswegen individuelle Entscheidungen bei der Techniknutzung ermöglicht werden sollten.

Zu den relevanten **rechtlichen Rahmenbedingungen** zählen u.a. die universellen **Menschenrechte**, (nationale) **Produkthaftungsgesetze**, die **KI-Verordnung** der Europäischen Union sowie die **Datenschutzgrundverordnung** der Europäischen Union (DSGVO). Letztere regelt den **datenschutzsensiblen** Umgang mit personenbezogenen Daten – ein Aspekt, der im Kontext häuslicher Assistenzsysteme besonders relevant ist. Dies entspricht der ethischen Grundanforderung der transparenten, nachvollziehbaren und freiwilligen Datenverarbeitung, die die Privatsphäre der Nutzenden respektiert und schützt.

[3] Für manche wäre eine 24-Stunden-Videoüberwachung zum Ziele der Sturzerkennung, Sicherheit und Autonomie akzeptabel, für andere ist ein solcher Eingriff in die Privatsphäre nicht tragbar und bedeutet in Bezug auf Autonomie/Unabhängigkeit genau das Gegenteil: sich nicht von Technologien abhängig zu machen.

Ethische Überlegungen – beispielsweise zur Frage, inwiefern Menschen Technologien selbstbestimmt nutzen können, die ihre Privatsphäre womöglich einschränken – gehen jedoch über rein rechtliche Fragestellungen hinaus. Im Folgenden werden exemplarisch ethikbezogene Kriterien und Dimensionen aus den Feldern der **Politik**, **Technik** und **Forschung** vorgestellt. Sie beziehen sich u.a. auf Ethik im Kontext von KI, Robotik, technischer Entwicklung im Allgemeinen sowie speziell auf soziale Robotik für ältere Menschen.

Die in den einschlägigen (internationalen) Empfehlungen formulierten ethischen Grundsätze sind teils abstrakt formuliert und auf Projektebene nicht immer umsetzbar (Panico et al., 2020, S. 36081).

Nichtsdestotrotz enthalten derlei (techno-)politische Forderungen zentrale ethikbezogene Ansprüche:

> „Die Entwicklung und der Einsatz digitaler Technologien, so auch Künstlicher Intelligenz, müssen im Rahmen der *Menschenrechte* und der UN-Agenda 2030 für *Nachhaltige Entwicklung* erfolgen. Die Politik steht daher weltweit vor der dringenden Aufgabe, Governance-Systeme für KI zu entwickeln, die *die Wahrung der Grund- und Freiheitsrechte, den Schutz gesellschaftlicher Vielfalt, die Sicherung und Verteidigung von Demokratie und Rechtsstaat und den Abbau bestehender Asymmetrien, wie etwa im Zugang zu Ressourcen und Macht*, gewährleisten."
> (Fachbeirat Ethik der Künstlichen Intelligenz, 2023, hinzugefügte Hervorhebungen)

Ähnlich abstrakt formuliert die UNESCO (2023: S. 43–51) folgende **Werte** in Bezug auf KI-Entwicklung:

- „Achtung, Schutz und Förderung der Menschenrechte, der Grundfreiheiten und der Menschenwürde" (S. 43)
- „Günstige Entwicklung von Umwelt und Ökosystemen" (S. 46)
- „Gewährleistung von Vielfalt und Inklusivität" (S. 50)
- „Leben in friedlichen, gerechten und verbundenen Gesellschaften" (S. 51)

Zudem formuliert die UNESCO (2023) etwas konkretere, jedoch noch immer teils weitereichende **Prinzipien**:

- „Verhältnismäßigkeit und Schadensverhütung“ (S. 53)
- „Sicherheit und Schutz“ (S. 55)
- „Fairness und Nichtdiskriminierung“ (S. 57)
- „Nachhaltigkeit“ (S. 64)
- „Recht auf Privatsphäre und Datenschutz“ (S. 65)
- „Menschliche Aufsicht und Entscheidung“ (S. 69)
- „Transparenz und Erklärbarkeit“ (S. 71)
- „Verantwortlichkeit und Rechenschaftspflicht“ (S. 79)
- „Bewusstsein und Kompetenzen“ (S. 81)
- „Flexible Organisation und Zusammenarbeit unter Einbeziehung aller Interessengruppen“ (UNESCO, 2023, S. 53–83).

Nevejans (2016) analysiert in ihrer Studie im Auftrag der Europäischen Union rechtliche Anforderungen, die mit roboethischen Prinzipien übereinstimmen, d.h. rechtlich abgesicherte ethische Prinzipien, die zum Schutze der Menschen vor Schaden durch Roboter umgesetzt werden sollen:

- Schutz des Menschen vor Schaden durch Roboter
- Achtung der Ablehnung von Pflege durch einen Roboter
- Schutz der menschlichen Freiheit im Umgang mit Robotern
- Schutz vor Verletzungen der Privatsphäre durch Roboter
- Umgang mit personenbezogenen Daten, die von Robotern verarbeitet werden
- Schutz der Menschheit vor dem Risiko der Manipulation durch Roboter
- Vermeidung des Verlusts sozialer Bindungen
- Gleichberechtigter Zugang zu Fortschritten in der Robotik
- Einschränkung des menschlichen Zugangs zu Enhancement-Technologien[4]

(Nevejans, 2016, S. 20–26).

[4] D.h. Technologien zur Optimierung menschlicher Leistungsfähigkeit (hier durch Robotik).

Nicht nur die Politik bzw. supranationale Akteur*innen wie die EU oder UNESCO, sondern auch Technologiekonzerne stellen ethische Prinzipien auf. So hat die Robert Bosch GmbH (2020) einen „KI-Kodex“ veröffentlicht, zu dessen Wahrung sich das Unternehmen bei der Entwicklung und Nutzung von KI-Produkten verpflichtet:

- „Jedes KI-Produkt von Bosch soll unserem Leitmotiv Technik fürs Leben, das *Innovationsstreben* mit *gesellschaftlicher Verantwortung* verbindet, entsprechen.
- KI soll *nicht ohne eine menschliche Kontrollinstanz über den Menschen entscheiden*, sondern *dem Menschen als Werkzeug* dienen.
- Wir wollen *sichere*, *robuste* und *nachvollziehbare* KI-Produkte entwickeln.
- Vertrauen gehört zu den Grundwerten unseres Unternehmens. Wir haben den Anspruch, *vertrauenswürdige* KI-Produkte zu entwickeln.
- Bei der Entwicklung von KI-Produkten beachten wir rechtliche Vorgaben und orientieren uns an *ethischen Grundsätzen*.“

  (Robert Bosch GmbH 2020; hinzugefügte Hervorhebungen)

Die in der Forschungsliteratur vorgelegten Ethikkriterien weisen Parallelen zu manchen der bisher angeführten abstrakten *ethischen Grundsätzen* auf. Barton & Pöppelbuß (2022) fassen auf der Basis verschiedener Ethik-Modelle sechs zentrale Prinzipien für die ethische Nutzung von künstlicher Intelligenz zusammen: *Wohltätigkeit*, *Transparenz*, *Nicht-Boshaftigkeit*, *Autonomie*, *Gerechtigkeit* und *Datenschutz*, die jedoch allgemein genug formuliert sind, um auch für andere Technologien herangezogen werden zu können.

Für den Kontext von sog. altersgerechten Assistenzsystemen im Pflege- und Gesundheitskontext entwickeln Manzeschke et al. (2013) und Weber (2015, 2016) ein Modell zur ethischen Evaluation sozio-technischer Arrangements (Akronym: MEESTAR). Dabei definieren sie sieben ethische Dimensionen: *Teilhabe*, *Fürsorge*, *Sicherheit*, *Selbstbestimmung*, *Gerechtigkeit*, *Privatheit* und *Selbstverständnis*. Da Bewertungen auf individueller, organisationeller und gesellschaftlicher Ebene womöglich unterschiedlich ausfallen, bedarf es einer Priorisierung bzw. Aushandlung zwischen konkurrierenden Interessen und Werten.

Komplementär zu diesen ethischen Dimensionen legen die Autor*innen 15 umfassendere Leitlinien vor:

1. Selbstbestimmung
2. Eingeschränkte Selbstbestimmung/Autonomie
3. Teilhabe
4. Gerechtigkeit
5. Sicherheit
6. Privatheit
7. Datenschutz
8. Aufklärung & informationelle Selbstbestimmung
9. Haftung
10. Altersbilder
11. Vermeiden von Diskriminierung und Normierung
12. Anwendungsfreundlichkeit
13. Vertragsbestimmungen
14. Qualifizierung und Weiterbildung
15. Verantwortung & bestmögliche Unterstützung durch Technik.

(Manzeschke et al., 2013, S. 22–26)

Für den Bereich der *Sozialen Robotik für ältere Menschen* formulieren Autor*innen wie Körtner (2016) Ethikkriterien, die im Einklang mit den bereits genannten, allgemeinen (positiv formulierten) Prinzipien stehen, wie *Würde*, *Privatsphäre*, *Sicherheit*, *Datenschutz*, *Autonomie*, *Fairness*, *Fürsorge* und *Verantwortlichkeit*. Ähnlich wie die oben zitierten roboethischen Prinzipien von Nevejans (2016) verweisen Arbeiten in diesem Feld auch auf spezifische **Risiken** bzw. **ethische Bedenken** wie *Täuschung*, *Verletzung der Würde (vulnerabler Gruppen)*, *soziale Isolation*, *Einschränkung der Privatsphäre und fehlende Kontrolle über Datenverwendung bzw. -speicherung*, *Gefährdung der Sicherheit durch den Roboter* und *Nichtberücksichtigung der Vulnerabilität von älteren Menschen* (Körtner 2016). Auch Sharkey & Sharkey (2012) gehen auf *verringerten menschlichen Kontakt*, *Objektifizierung und Kontrollverlust*, *Einschränkung der Privatsphäre und persönlichen Freiheit*, *Täuschung und Infantilisierung* und *Fragen der Verantwortlichkeit* ein.

Aus den hier vorgestellten ethikbezogenen Kriterien und Prinzipien aus *Politik*, *Recht*, *Forschung* und *Wirtschaft* lassen sich die folgenden Gemeinsamkeiten zusammenfassen:

- Menschenzentrierung und Schutz der Grundrechte
- Verantwortung und Rechenschaftspflicht
- Transparenz und Nachvollziehbarkeit
- Schutz vor Schaden und Missbrauch
- Privatsphäre und Datenschutz
- Gerechtigkeit, Fairness und Inklusion
- Menschliche Kontrolle und Entscheidungsfreiheit
- Nachhaltigkeit und gesellschaftliche Verantwortung
- Kontextsensitivität und Aushandlung verschiedener Interessen

Aus den zitierten Quellen geht auch ein Konsens über eine verantwortungsvolle Technikgestaltung hervor, die durch die Berücksichtigung von ethischen Kriterien sichergestellt werden soll. Neben diesen allgemeinen Prinzipien verweist die Fachliteratur auch auf gewisse Spezifika, die sich für die **soziale Robotik** und insbesondere die **Zielgruppe älterer Menschen** ergeben. So vereinen ethikbezogene Kriterien im Bereich sozialer Robotik für ältere Menschen zentrale normative Prinzipien wie Würde, Autonomie und Datenschutz mit spezifischen Risiken wie Täuschung, Isolation oder Kontrollverlust. Zwar werden hier potenziell zielgruppenspezifische Vorsichtsmaßnahmen gesetzt, was v.a. in der Formulierung von ethischen **Risiken** zum Ausdruck kommt: Diese Zielgruppe wird als potenziell vulnerabel angesehen und technologische Eingriffe können tief in ihre Privatsphäre, Autonomie und sozialen Beziehungen eingreifen. Daher ist besondere Vorsicht geboten, um Überforderung, Entmündigung oder unbeabsichtigte soziale Isolation zu vermeiden. Um jedoch stereotype Altersbilder zu vermeiden (Manzeschke et al., 2013, S. 24), wird in dieser Arbeit argumentiert, die Nutzer*innenperspektive und den Anwendungskontext konsequent in der Entwicklung zu berücksichtigen, jedoch nicht im Vorhinein von bestimmten altersspezifischen Eigenschaften oder Bedürfnissen auszugehen. Diese müssen erst empirisch erhoben werden. Folglich können die im nächsten Abschnitt vorgestellten Ethikkriterien auch als **allgemeine Ethikkriterien für die Entwicklung von technischen Assistenzsystemen** verstanden werden.

4 Ethikkriterien: Definition und praktische Umsetzung

Welche Ethikkriterien erweisen sich nun als relevant für die Entwicklung eines sprechenden Staubsaugerroboters zur Sturzerkennung für ältere Menschen? Im Rahmen des Vorgängerprojekts (**Smart Companion**) wurden **16 Ethikkriterien**[5] definiert, die sich u.a. an MEESTAR (Manzeschke et al., 2013; Weber, 2015, 2016; siehe letzter Abschnitt) und weiteren bereits vorgestellten roboterspezifischen Ethikdimensionen orientieren und auf die Entwicklung altersgerechter Assistenzsysteme wie Smart Companion zugeschnitten sind:

1. Vorbeugung gegen Täuschung
2. Sicherheit der Nutzer*innen
3. Privatsphäre
4. Datensicherheit
5. Fairer Umgang mit vulnerablen Personen
6. Achtung der Würde der Nutzer*innen
7. Transparenz & Nachvollziehbarkeit
8. Autonomieförderung
9. Inklusion & Solidarität
10. Lebensqualität fördern
11. Gleichheitsprinzip
12. Nicht-Unterwerfung
13. Verbandliche Risikoverantwortlichkeit
14. Nutzer*innenpartizipation in der Entwicklung und Konfiguration
15. Ersatz von humaner Arbeitskraft
16. Nachhaltigkeit in der Entwicklung

Diese 16 Ethikkriterien wurden im Laufe des Folgeprojekts **Smart Companion 2** überarbeitet und reduziert.

[5] Hier sei u.a. auf die federführende Rolle und Vorarbeiten von Florian Zahorka hingewiesen. Siehe auch https://research.fhstp.ac.at/content/download/167912/file/Ethik_Kriterien_Beschreibung.pdf

KAPITEL 4 – ETHIKKRITERIEN – DEFINITION UND PRAKTISCHE UMSETZUNG

Die nun **sieben Ethikkriterien**[6] gehen weiters u.a. auf Forschungsliteratur (siehe letzter Abschnitt), Workshops mit Projektpartner*innen, dem Entwicklungsteam, dem projektinternen Ethikboard und den Ausarbeitungen für die Einreichung des Ethikantrags bei der Ethikkommission der Medizinischen Fakultät der Johannes Kepler Universität Linz für die Durchführung der Pilotfeldstudie in Linz[7] zurück.

Im Laufe des Erarbeitungsprozesses waren wir ebenfalls mit der oben erwähnten Herausforderung der Aushandlung potenziell konfligierender Interessen und Bedürfnisse konfrontiert. Dies zeigte sich unter anderem in den für Hausnotrufsysteme typischen Spannungsfeldern: dem Bedürfnis von Angehörigen nach engmaschiger Kontakt- und Betreuungsmöglichkeit, der Sorge der Primärnutzer*innen vor Überwachung sowie dem hohen Ressourcenaufwand für Heimnotruforganisationen infolge einer erhöhten Falsch-Positiv-Rate[8].

Im Folgenden wird jeweils das Ethikkriterium allgemein für den Kontext von Assistenzsystemen vorgestellt und dann am Beispiel von Smart Companion anhand von ausgewählten Dimensionen konkretisiert. Vorausgeschickt sei zudem, dass es weiterführende ethikbezogene Prinzipien gibt, auf die im Folgenden nicht näher eingegangen wird. Dies geht u.a. drauf zurück, dass lediglich jene Kriterien herangezogen wurden, die operationalisierbar, d.h. konkret fassbar und in konkreten Designentscheidungen anwendbar sind.

[6] Die Ethikkriterien wurden vom Ethikboard bestehend aus Daniel Green, Andreas Jakl, Michael Kaindleinsberger, Sabine Lehner, Birgit Meinhard-Schiebel, Angelika Neuer, Johannes Pflegerl, Veronika Schauer & Henrik Smith (alphabetische Reihenfolge) kollaborativ geschrieben. Sabine Lehner hat den kollaborativen Schreibprozess angeleitet und die Überarbeitung letztverantwortet.

[7] Positives Votum der Ethikkommission der Medizinischen Fakultät Johannes Kepler Universität Linz (EK-Nr: 1022/2024).

[8] Das Auftreten von Dilemmata bei Entwicklungsentscheidungen sei am folgenden Beispiel veranschaulicht: Während ursprünglich angedacht war, dass Nutzer*innen entscheiden können, wer kontaktiert werden soll – beispielsweise eine Notruforganisation oder eine Zu/Angehörige, und dies auf große Zustimmung bei den Primärnutzer*innen stieß, wurde dies im Zuge der Diskussionen mit verschiedenen Stakeholder*innen wieder verworfen. So überwog für uns die Überlegung, dass eine potenzielle Notsituation von geschulten Heimnotruf-Disponent*innen besser abgeklärt werden kann als von einer Privatperson. Zudem sollte vermieden werden, dass es bei einem möglichen Notfall bspw. zu einem Nichtabheben oder einer Überforderung kommt. Letztlich würden sich in diesem Fall auch komplexere rechtliche Fragen stellen, die bei Heimnotruforganisationen aufgrund ihrer Erfahrungen mit Heimnotrufsystemen bereits geklärt sind.

Zu den nicht näher ausformulierten Prinzipien zählt erstens *die Wahrung der Menschenrechte*[9] und *demokratischer Prinzipien*, was als grundlegend vorausgesetzt und ausnahmslos eingehalten wird. Zweitens gibt es Prinzipien, die zwar als relevant erachtet werden, jedoch außerhalb des Einflussbereichs bzw. Realisierungsmöglichkeiten von einzelnen Projekten liegen.

Dazu zählen beispielsweise die Orientierung an *Nachhaltigkeitsprinzipien* (im ökologischen Sinne die Produktionsbedingungen betreffend), die *Förderung von gesellschaftlicher Verbundenheit durch Technologien* oder *Haftungsfragen*. Daher beschränken sich die folgenden ausgewählten Ethikkriterien auf jene Dimensionen, die unmittelbar in Designentscheidungen bspw. im Rahmen eines Forschungs- und Entwicklungsprojekts berücksichtigt werden können (Nussbaumer et al., 2021 für eine ähnliche ethikbasierte Entwicklung in einem konkreten Projekt).

Kriterium 1: Sicherheit der Nutzenden

Allgemein: Eingesetzte Technologien sind am neuesten Stand der Technik und gewährleisten die Betriebssicherheit geltender Normen. Es sind technische Maßnahmen zu implementieren, die die Funktionalität des Geräts überwachen und Fehlfunktionen erkennen und anzeigen. Dies umfasst auch die Implementierung eines Sicherheitsmechanismus, der eine sofortige Deaktivierung des Systems im Falle einer gravierenden Fehlfunktion veranlasst.

Umsetzung im Smart Companion: Da der Smart Companion als mobiler Staubsaugerroboter durch die Wohnung fährt, sind Maßnahmen wie akustische oder visuelle Signale erforderlich, um das von ihm ausgehende Stolperrisiko zu minimieren. Der Staubsaugerroboter kann verschiedene Hindernisse erkennen und ihnen nach Möglichkeit ausweichen.

Zur Sicherstellung eines zuverlässigen Betriebs wurde eine sogenannte State Machine (Zustandsmaschine) für die Steuerung des Smart Companion entwickelt. Diese legt das Systemverhalten in erwartbaren Situationen fest und sorgt dafür, dass die Auswirkungen fehlerhafter Komponenten kontrolliert bleiben.

[9] Wobei manche Ethikkriterien durchaus durch die Menschenrechte abgedeckt sind wie Würde, Selbstbestimmung, Freiheit und Sicherheit (Brey & Dainow, 2024, S. 1268).

Im Falle einer Störung stellt die Steuerung sicher, dass das System in keinen undefinierten Zustand gerät. Dadurch werden Fehler-Folgen minimiert und ein zuverlässiger Betrieb aufrechterhalten.

Weiters wird die sichere Funktionalität durch die automatisierte Überprüfung mehrerer Parameter, sowie Erkennung und Anzeige von Fehlfunktionen, gewährleistet. Fehlermeldungen können ggf. an ein betreuendes, technisches Team zur Fehlerbehebung übermittelt werden. Fehlermeldungen werden u.a. angezeigt, wenn sich der Roboter verfangen hat, oder die Kommunikation zwischen den technischen Komponenten nicht ordnungsgemäß funktioniert. Es werden Handlungspfade definiert, die im Falle von Fehlfunktionen oder anderen Gründen, welche die Betriebssicherheit gefährden, greifen. Dies umfasst beispielsweise den Abbruch von Reinigungsfahrten bei niedrigem Akkustand, oder eine Temperaturüberwachung des Akkus.

Neben der Möglichkeit, sämtliche Handlungen durch verbale Befehle zu unterbrechen, ermöglicht ein Notausschalter, Smart Companion sofort zu deaktivieren.

Einschulungen und Anleitungen (in verschiedenen Formen und Sprachen) tragen zur sicheren Handhabe bei. Zusätzlich werden die Nutzenden aufgeklärt, an wen sie sich bei Fragen und Fehlfunktionen wenden können (bspw. eine Service-Hotline).

Die sichere Funktionalität wurde durch mehrere Testverfahren (im Laborsetting und sechs einmonatige Probenutzungen) überprüft und entsprechende technische Dokumentationen wurden angefertigt. Für ein serienreifes Produkt wird eine externe Zertifizierung bzgl. Produktsicherheit empfohlen.

Kriterium 2: Datenschutz und Privatsphäre

Allgemein: Der Schutz der Privatsphäre wird durch transparente Datenschutzrichtlinien, Verschlüsselung und eine datensparsame Gestaltung von Technologien gewährleistet. Dies umfasst auch das Recht und die Kontrolle einer Person über ihre persönlichen Daten, ihre digitale Selbstbestimmung und den Schutz vor ungewollter Überwachung oder Datenmissbrauch.

Es werden somit lediglich jene Daten erhoben, die für eine sichere Funktionalität und dem Zweck entsprechend notwendig sind. Relevante Rechtsgrundlagen und nationale Bestimmungen werden eingehalten.

Die Nutzenden haben die Möglichkeit, in der Konfiguration Entscheidungen über die Produktnutzung und damit verbundene Datennutzung in Form von verschiedenen Zustimmungs-Abstufungen zu treffen.

Es wird aufgeklärt,

- welche Daten (Datenminimierung),
- zu welchem Zweck (Zweckbestimmung),
- wo, von wem, wann,
- wie lange erhoben, gespeichert, verarbeitet,
- für Nutzende zugänglich (Transparenz) gemacht
- und gelöscht werden (können).

Sofern Aufzeichnungen vorgenommen werden sollen, muss die Möglichkeit bestehen, die Aufzeichnung auch abzulehnen (informationelle Selbstbestimmung). Wenn dadurch die Funktionalitäten eingeschränkt oder nicht erfüllt werden, müssen die Nutzenden darüber aufgeklärt werden. Eine Nutzung der Daten ist lediglich im Rahmen der jeweils relevanten Rechtslagen (bspw. Erfüllung eines Vertrages) möglich.

Im Rahmen einer Datenschutzfolgeabschätzung sind Maßnahmen zur Minimierung von Risiken zu setzen.

Bei datenschutzrechtlichen Angelegenheiten handelt es sich um komplexe Inhalte, für deren Zustimmung es eine verständliche (sprachliche) Form sowie ggf. entsprechende Unterstützung benötigt. Es wird empfohlen, die jeweiligen Zustimmungs-Optionen übersichtlich zu gestalten.

Umsetzung im Smart Companion: Der Umgang mit den Daten (Erhebung, Speicherung, Übermittlung, Löschung etc.) ist konform mit der Datenschutz-Grundverordnung der Europäischen Union (DSGVO), wird den Nutzenden transparent gemacht und basiert ausschließlich auf ihrer expliziten Zustimmung.

Dies betrifft beispielsweise die Erhebung und Analyse von Fotos, die für die Hindernis- und Sturzerkennung genutzt werden. Diese Fotos, die der Staubsaugerroboter laufend zur eigenen Navigation anfertigt, werden mittels Bildanalyse vom Smart Companion ausgewertet. Hervorzuheben ist, dass die Hindernis- und Sturzerkennungsfunktion ohne Zustimmung zur Fotoaufnahme technisch nicht möglich ist. Bei Nicht-Zustimmung der Bildanalyse umfasst der Smart Companion lediglich eine Staubsaugfunktion. Smart Companion kann dennoch jederzeit verbal um Hilfe gebeten werden.

Im Sinne der Datenminimierung und zum Schutze der Privatsphäre werden sämtliche Fotos bis auf den Ausnahmefall eines vermuteten Sturzes, nach dem der*die Nutzende nicht mehr ansprechbar ist, automatisch gelöscht. Die Fotos sind für andere nicht einsehbar. Somit werden standardmäßig keine Fotos dauerhaft gespeichert oder übermittelt.

Sofern Smart Companion mit einer Notruforganisation verbunden ist, können die Nutzenden festlegen, ob im Falle eines Sturzes ein Foto über gesicherte Kanäle verschlüsselt an die Notruforganisation übertragen wird. Damit können sich die Notrufstellenmitarbeitenden leichter ein Bild über die Lage machen. Nach der Übertragung werden die Fotos automatisch gelöscht, sodass keine Fotos im Roboter gespeichert sind. Bei einer fehlenden Zustimmung wird kein Foto übertragen und lediglich die Notruforganisation kontaktiert.

Zur Wahrung der Privatsphäre (z.B. im Badezimmer) kann die Bewegung des Smart Companion durch die Festlegung von sog. No-Go Zonen, d.h. Zonen, in die Smart Companion nicht fahren darf, eingeschränkt werden. Auch hier sind die Nutzenden darüber aufzuklären, dass der Smart Companion in diesen Zonen seine Sturzerkennungsfunktion nicht erfüllen kann. Es ist jedoch nicht auszuschließen, dass Smart Companion in diesen Zonen nicht doch Fotos beim Vorbeifahren aufnimmt.

Kriterium 3: Transparenz und Nachvollziehbarkeit

Allgemein: Die Funktionsweise des Geräts wird transparent und nachvollziehbar vermittelt. Nutzende können dadurch verstehen, wie sich das Gerät verhält und wie Entscheidungen zustande kommen.

Leicht zugängliche Informationsmaterialien, welche die Möglichkeiten und Grenzen des Assistenzsystems verständlich erläutern, werden bereitgestellt.

Umsetzung im Smart Companion: Die jeweiligen Schritte des Vorgehens werden nachvollziehbar dargestellt. So fährt Smart Companion nach einem festgelegten Fahrtenplan oder gemäß den Anweisungen der Nutzenden. Smart Companion kündigt seine Handlungen mittels Sprachausgabe an. Weitere Schritte werden bei den Nutzenden erfragt (z.B. Initiierung eines Anrufs beim Notrufdienst).

Die Nutzenden werden in einer Einschulung über die Funktionsweise und der Umgang mit den Daten (siehe Punkt 4.2) aufgeklärt. Es werden zusätzliche Materialien wie ein Übersichtsblatt mit den wichtigen Funktionsweisen und Befehlen und Videos zur Verfügung gestellt. Die Informationsmaterialien beschreiben zudem die Szenarien und legen dar, wie Entscheidungen in Notfallszenarien getroffen werden.

Kriterium 4: Achtung der Würde der Nutzenden

Allgemein: Ein Assistenzsystem muss die Würde des Menschen achten, indem sichergestellt wird, dass Nutzende die Kontrolle über das System behalten und nicht entmündigt oder bevormundet werden. Bei der Entwicklung und Nutzung des Assistenzsystems ist darauf zu achten, dass keine Klassifizierung von menschlichen Eigenschaften oder medizinischen Diagnosen möglich sind (siehe KI-Gesetz der Europäischen Union[10]). Wesentlich ist auch die Reflexion und Kommunikation mit involvierten Personengruppen darüber, für welche Personen und in welchen Zusammenhängen die Technologie geeignet ist. Das betrifft auch den Vertrieb und die Bewerbung des Assistenzsystems.

Die Würde der Nutzenden wird auch durch die Prinzipien des Datenschutzes (Punkt 4.2) und der Transparenz (Punkt 4.3) gewahrt.

Umsetzung im Smart Companion: Smart Companion ist kein Gerät zur Überwachung und Kontrolle von Menschen, sondern dient zum Erkennen und zur Unterstützung in Sturzsituationen. Um die Würde der Nutzenden zu wahren, können diese auswählen, in welche Räume/Zonen der Smart Companion fahren darf.

[10] Siehe (Europäische Union, 2024).

Eine Foto-Übertragung an die Notruforganisation findet nur bei einem vermuteten ‚Notfall' – sofern dies in der Konfiguration so eingestellt wurde – statt (siehe 4.3).

Bei der Bewerbung wird auf altersdiskriminierende und sonstige stereotype Repräsentationen verzichtet. In der Sprachinteraktion wird darauf geachtet, dass keine diskriminierenden Formulierungen gewählt werden. Das Bedürfnis nach sozialen Kontakten, die andere Hilfsangebote erfüllen können, kann durch Smart Companion nicht substituiert werden.

Kriterium 5: Autonomie und Nicht-Unterwerfung

Allgemein: Der eigene Wille von Nutzenden wird berücksichtigt. Assistenzsysteme fällen keine eigenen, selbst hergeleiteten Entscheidungen, sondern operieren lediglich aufgrund vorab definierter Entscheidungsstrukturen. Die Letztentscheidung bei kritischen Vorgängen bleibt dabei stets beim Menschen.

Nutzende haben die Möglichkeit, in der Konfiguration Entscheidungen über die Produktnutzung und weitere Dimensionen (wie Datenschutz) zu treffen (siehe auch 4.3) und Assistenzsysteme mittels entsprechender Anweisungen anlassbezogen zu steuern.

Umsetzung im Smart Companion: Sowohl in der Konfiguration als auch in der Anwendung haben Nutzende Auswahlmöglichkeiten bspw. in Bezug auf den Reinigungs- und Suchfahrtenzeitplan. Sie können zusätzlich angeben, dass Smart Companion am heutigen Tag nicht fahren soll, beispielsweise weil Besuch erwartet wird. Wenn vermutet wird, dass der*die Bewohner*in gestürzt ist, wird mittels Sprachassistenz geklärt, ob Hilfe gerufen werden soll. Handlungen von Smart Companion können jederzeit unterbrochen werden (verbal oder durch einen Schalter).

Kriterium 6: Vermeidung von Diskriminierung und Achtung von Diversität

Allgemein: Die Technologie darf nicht zu einer Diskriminierung oder zu einem weiteren Ausschluss von potenziell vulnerablen Personengruppen führen. Es ist auf die Diversität von Menschen, insbesondere auch ihre spezifischen Bedürfnisse Rücksicht zu nehmen. Eine Diskriminierung aufgrund ethnischer oder nationaler Zugehörigkeit, Religion, Weltanschauung, Behinderung, Alter, sexueller Orientierung, Geschlecht bzw.

Geschlechtsidentität, Sprache o. Ä. ist zu verhindern. Zudem sind weitere Faktoren und Umstände wie geringe Erfahrung mit Technologien in der Entwicklung zu berücksichtigen.

Der Zugang zu Information, Wissen und Ressourcen soll für alle Nutzenden gleichermaßen ermöglicht werden.

Umsetzung im Smart Companion: Das Training der künstlichen Intelligenz zur Identifizierung von gestürzten Personen inkludiert verschiedene Diversitätsmerkmale (wie Hautfarbe, Alter, Geschlecht, Behinderung, sonstige körperliche Merkmale, verschiedene Zielgruppen), um die korrekte Erkennung zu gewährleisten sowie einen Bias zu reduzieren.

Die Dialoge sind in leichter Sprache gehalten und die Anpassung der Lautstärke ist möglich. Zusätzliche Sprachmodule für verschiedene Sprachen sind vorhanden und umfassen auch verschiedene Dialekte.

Außerdem wurde darauf geachtet, dass die eingebaute Sprachassistenz ausreichend Zeit zum Antworten gibt, was für Menschen mit verlangsamter Sprachproduktion, Stottern oder anderen Einschränkungen wichtig ist. Dies soll eine einfache, unmissverständliche und zuverlässige Kommunikation im Alltag und in Sturzsituationen ermöglichen. Auch das Informations- und Erklärungsmaterial ist in verständlicher Sprache verfasst und steht in multimodalen Formen zur Verfügung.

Es ist auf den Umstand zu achten, dass die Inanspruchnahme von Assistenzsystemen potenziell mit negativen Erfahrungen bzw. Stigmatisierung assoziiert wird (bspw. in Bezug auf mögliche Abhängigkeitsgefühle).

Die Zielgruppe ist klar zu definieren. Für Smart Companion sind damit sturzgefährdete und tendenziell ältere Personen gemeint, die in der Lage sind, die Funktionsweise von Smart Companion zu verstehen, ihre Zustimmung zur Nutzung und Datenverarbeitung zu geben und ihren eigenen Willen auszudrücken. Dies bedeutet jedoch auch, dass der Smart Companion für gewisse Personengruppen nicht geeignet ist, wie zum Beispiel Personen mit einer (fortgeschrittenen) Demenzerkrankung oder gehörlose Personen. Weitere Zielgruppen wie jüngere, tendenziell nicht sturzgefährdete Menschen, können von Smart Companion

ebenfalls profitieren, werden aber in Bezug auf Klarheit in der Produktentwicklung sowie der damit einhergehenden Entwicklung nicht explizit angesprochen.

Kriterium 7: Vorbeugung gegen Täuschung und Verwechslung

Allgemein: Das Assistenzsystem ist so zu gestalten, dass keine Verwechslung mit einem Menschen bzw. einem denkenden und fühlenden Wesen möglich ist. Daher sind jene Eigenschaften zu vermeiden, die den Eindruck von menschlichen Fähigkeiten wie Empathie, soziale Nähe oder Gefühle suggerieren. Dadurch lässt sich auch eine etwaige Fehleinschätzung der Fähigkeiten des Assistenzsystems vermeiden. Jegliche Irreführung der Nutzenden ist unzulässig.

Umsetzung im Smart Companion: Um eine Verwechslung mit einem Menschen und – damit einhergehend – ein falsches Sicherheitsgefühl zu vermeiden, erfolgt eine persönliche Aufklärung über die Möglichkeiten und Grenzen von Smart Companion.

Dies bedeutet, die Nutzenden aufzuklären, dass hinter dieser Anwendung kein Mensch steckt und künstliche Intelligenz nur für spezifische Aufgaben wie die Bilderkennung und Spracherkennung eingesetzt wird. Smart Companion umfasst nur auf definierte Nutzungsszenarien begrenzte Interaktionen. So ist es im Rahmen der Anwendung möglich, dass die Person gefragt wird, ob Hilfe benötigt wird oder sie einen Hilferuf durch den Smart Companion initiiert.

Es ist jedoch anzumerken, dass bei der Testung des Smart Companion Prototypen (in der sog. Pilotfeldstudie) manche Teilnehmenden gerne die zusätzlichen Funktionen des Sprachassistenz-Systems (Amazon Alexa) genutzt haben, um zum Beispiel Musik zu hören. Diese zusätzlichen Unterhaltungszwecke erhöhen die Interaktionsfrequenz und Gefahr, dass Smart Companion als Interaktionspartner wahrgenommen wird. Diese erweiterte Funktionalität geht jedoch über das Kernprodukt der Smart Companion Funktionen hinaus. Angesichts dieser gut angenommenen Funktion auf der einen Seite und der Täuschungs- und Vermenschlichungsvermeidung auf der anderen Seite wäre eine mögliche Erweiterung der Funktionalität in einem weiterführend entwickelten Produkt zu klären.

Literatur

Barton, M.-C., & Pöppelbuß, J. (2022). Prinzipien für die ethische Nutzung künstlicher Intelligenz. HMD Praxis der Wirtschaftsinformatik, 59(2), 468–481. https://doi.org/10.1365/s40702-022-00850-3

Brey, P., & Dainow, B. (2024). Ethics by design for artificial intelligence. AI and Ethics, 4(4), 1265–1277. https://doi.org/10.1007/s43681-023-00330-4

Cavoukian, A. (2011). Privacy by Design in Law, Policy, and Practice,. Canadian Electronic Library. https://coilink.org/20.500.12592/ccgc2f

Coughlin, J. F., D'Ambrosio, L. A., Reimer, B., & Pratt, M. R. (2007). Older Adult Perceptions of Smart Home Technologies: Implications for Research, Policy Market Innovations in Healthcare. 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1810–1815. https://doi.org/10.1109/IEMBS.2007.4352665

Europäische Union. (2024). Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz)Text von Bedeutung für den EWR.

Fachbeirat Ethik der Künstlichen Intelligenz. (2023). Zur Dringlichkeit der Umsetzung der UNESCO-Empfehlung zur Ethik der Künstlichen Intelligenz. Stellungnahme des Fachbeirats Ethik der Künstlichen Intelligenz der Österreichischen UNESCO-Kommission anlässlich seiner Konstituierung. https://www.unesco.at/fileadmin/user_upload/Stellungnahme_des_Fachbeirats_Ethik_der_Kuenstlichen_Intelligenz_05-07-2023.pdf

Heinbüchner, B., Hautzinger, M., Becker, C., & Pfeiffer, K. (2010). Satisfaction and use of personal emergency response systems. Zeitschrift Für Gerontologie Und Geriatrie, 43(4), 219–223. https://doi.org/10.1007/s00391-010-0127-4

Huang, T., & Huang, C. (2020). Elderly's acceptance of companion robots from the perspective of user factors. Universal Access in the Information Society, 19(4), 935–948. https://doi.org/10.1007/s10209-019-00692-9

Körtner, T. (2016). Ethical challenges in the use of social service robots for elderly people. Zeitschrift Für Gerontologie Und Geriatrie, 49(4), 303–307. https://doi.org/10.1007/s00391-016-1066-5

Manzeschke, A., Weber, K., Rother, E., & Fangerau, H. (Hrsg.). (2013). Ergebnisse der Studie „Ethische Fragen im Bereich Altersgerechter Assistenzsysteme". VDI.

Nevejans, N. (2016). European Civil Law Rules in Robotics. Study for the JURI Comittee. https://www.europarl.europa.eu/RegData/etudes/STUD/2016/571379/IPOL_STU(2016)571379_EN.pdf

Nussbaumer, A., Pope, A., & Neville, K. (2021). A framework for applying ETHICS-BY-DESIGN to decision support systems for emergency management. Information Systems Journal, 33(1), 34–55. https://doi.org/10.1111/isj.12350

Panico, F., Cordasco, G., Vogel, C., Trojano, L., & Esposito, A. (2020). Ethical issues in assistive ambient living technologies for ageing well. Multimedia Tools and Applications, 79(47–48), 36077–36089. https://doi.org/10.1007/s11042-020-09313-7

Robert Bosch GmbH. (2020, Februar 19). KI-Kodex von Bosch im Überblick. https://assets.bosch.com/media/de/global/stories/ai_codex/bosch-code-of-ethics-for-ai.pdf

Sharkey, A., & Sharkey, N. (2012). Granny and the robots: Ethical issues in robot care for the elderly. Ethics and Information Technology, 14(1), 27–40. https://doi.org/10.1007/s10676-010-9234-6

UNESCO. (2023). UNESCO-Empfehlung zur Ethik der Künstlichen Intelligenz. Deutsche UNESCO-Kommission e.V. https://www.unesco.at/fileadmin/Redaktion/Wissenschaft/Wissenschafts-_und_Bioethik/UNESCO-Empfehlung_zur_Ethik_der_Kuenstlichen_Intelligenz.pdf

Weber, K. (2015). MEESTAR: Ein Modell zur ethischen Evaluierung sozio-technischer Arrangements in der Pflege- und Gesundheitsversorgung (S. 247–262).

Weber, K. (2016). Ein erweitertes Modell zur ethischen Evaluierung soziotechnischer Arrangements. Technische Unterstützungssysteme.

Werthner, H. et al. 2019. Wiener Manifest für digitalen Humanismus. Verfügbar unter: https://caiml.org/dighum/dighum-manifesto/Vienna_Manifesto_on_Digital_Humanism_DE.pdf [14.07.2025].