

Ergebnisbericht

(gemäß Nr. 14.1 ANBest-IF)



Konsortialführung:	Fraunhofer-Gesellschaft für angewandte Forschung e.V.
Förderkennzeichen:	01VSF20016
Akronym:	KI-BA
Projekttitel:	Künstliche Intelligenz in der Versorgung-Bedingung der Akzeptanz von Versicherten
Autoren:	Dr. Stefan Kamin, Martina Simon
Förderzeitraum:	1. August 2021 – 31. Januar 2024

Inhaltsverzeichnis

I.	Abkürzungsverzeichnis	3
II.	Abbildungsverzeichnis	3
III.	Tabellenverzeichnis	4
1.	Zusammenfassung	5
2.	Beteiligte Projektpartner	6
3.	Projektziele	6
4.	Projektdurchführung	7
5.	Methodik.....	10
5.1	Methodik Kategorisierung	10
5.2	Methodik Akzeptanzstudie	14
5.3	Methodik Handlungsempfehlungen.....	20
6.	Projektergebnisse	20
6.1	Projektergebnisse Kategorisierung.....	20
6.2	Projektergebnisse Akzeptanzstudie	25
6.2.	Studie 1: Akzeptanz gegenüber KI-Anwendungsbeispielen	26
6.2.2	Studie 2: KI-Akzeptanz-Skala	32
6.2.3	Studie 3: Prädiktoren der KI-Akzeptanz.....	41
6.3	Projektergebnisse Handlungsempfehlungen	43
7.	Diskussion der Projektergebnisse.....	46
7.1	Kategorisierung.....	46
7.2	Akzeptanzstudie.....	47
7.3	Handlungsempfehlungen	47
8.	Verwendung der Ergebnisse nach Ende der Förderung.....	47

Akronym: KI-BA

Förderkennzeichen: 01VSF20016

9.	Erfolgte bzw. geplante Veröffentlichungen.....	49
10.	Literaturverzeichnis	49
11.	Anhang.....	50
12.	Anlagen	50

I. Abkürzungsverzeichnis

AP	Arbeitspaket
ATAI	Attitude Towards Artificial Intelligence
CAPI	Persönlich-mündliche Befragung
CAWI	Online-Fragebogen
df	Freiheitsgrade
E	Fehlerwahrscheinlichkeit
EFA	Explorative Faktorenanalyse
f	Effektstärke
GKV	Gesetzliche Krankenversicherung
KI	Künstliche Intelligenz
M	Mittelwert
MINRES	Minimum-Residuen-Methode
N/n	Größe der Stichprobe
p	p-Wert
PMM	Predictive Mean Matching
r	Pearson-Korrelationskoeffizient
RMSEA	Root Mean Square Error of Approximation
SD	Standardabweichung
SRMR	Standardized Root Mean Square Residual
T	t-Wert
TAM	Technology Acceptance Model
UV	Unabhängige Variable
Z	Konfidenzniveau

II. Abbildungsverzeichnis

Abbildung 1: Schematischer Überblick über den Projektablauf	8
Abbildung 2: Übersicht über das wissenschaftliche und methodische Vorgehen	10
Abbildung 3: Systematische Literaturanalyse zur Erstellung eines ersten Kategorisierungsansatzes	12
Abbildung 4: Aufbau der Befragung	19
Abbildung 5: Erster Entwurf eines Kategorisierungssystems	21
Abbildung 6: Übersicht aller KI-Beispiele aus den Experteninterviews und Einordnung ins Kategorisierungssystem	22
Abbildung 7: Finales Kategorisierungssystem für KI im Gesundheitswesen	23
Abbildung 8: Ausschnitt aus dem Videostimulus "Bildbasiertes Erkennen von Darmpolypen" - links Pretest, rechts finale Version	24

Abbildung 9: Haupteffekte der experimentellen Faktoren auf die KI-Akzeptanz.....	29
Abbildung 10: Interaktionseffekt KI-Erklärbarkeit und KI-Anwendung	30
Abbildung 11: Interaktionseffekt KI-Erklärbarkeit und Gruppe.....	31
Abbildung 12: Dreifach-Interaktion KI-Erklärbarkeit, KI-Autonomie und Gruppe.....	32
Abbildung 13: Strukturgleichungsmodell zur Messung von KI-Akzeptanz	40
Abbildung 14: Zentrale Anforderungen im Rahmen der konsolidierenden Abschlussdiskussion	46

III. Tabellenverzeichnis

Tabelle 1: Beteiligte Projektpartner und Verantwortlichkeiten	6
Tabelle 2: Zeitlicher Ablauf des Projekts inkl. Zuständigkeiten	9
Tabelle 3: Ein- und Ausschlusskriterien für die Literatursuche.....	11
Tabelle 4: Bei Versicherten und ÄrztInnen erfasste Konstrukte.....	17
Tabelle 5: Nur bei Versicherten erfasste Konstrukte	18
Tabelle 6: Nur bei ÄrztInnen erfasste Konstrukte.....	18
Tabelle 7: Einschätzung des Kategorisierungssystems	22
Tabelle 8: Übersicht über Stimulusvideos.....	24
Tabelle 9: Beschreibung der Stichprobe	27
Tabelle 10: Ergebnisse der ANOVA	28
Tabelle 11: Deskription	36
Tabelle 12: Darstellung der Ladungsstruktur	38
Tabelle 13: Handlungsempfehlungen zur Entwicklung und Implementierung von KI.....	44

1. Zusammenfassung

Die digitale Transformation und die damit verbundenen Möglichkeiten, klinische Daten in immer besserer Qualität und großer Menge zu erfassen, zu übertragen und zu verarbeiten, rückt zunehmend die Potenziale Künstlicher Intelligenz (KI) für die Versorgung im Rahmen der Gesetzlichen Krankenversicherung (GKV) in den Vordergrund. Jedoch besteht auf Seiten von Patienten und ÄrztInnen gleichermaßen häufig große Unsicherheit, wofür KI im Gesundheitswesen eingesetzt werden kann, welche Risiken und Potenziale bestehen und die die Qualität und Transparenz von KI-Anwendungen sichergestellt werden. Dabei kommt der Akzeptanz auf Seiten potenzieller Nutzer eine entscheidende Rolle für eine erfolgreiche Implementierung von KI-Lösungen zu (Stöber, Williger, Meerkamm, & Lang, 2012).

Vor diesem Hintergrund untersuchte das Projekt »KI-BA« die Akzeptanz von KI-Anwendungen in beispielhaften Versorgungsbereichen der GKV. Dabei wurden die nachfolgenden Projektziele adressiert:

- Projektziel 1: Entwicklung eines mehrdimensionalen Kategorisierungsinstrumentes für KI-Anwendungen in der medizinischen Versorgung unter Anwendung eines Mixed-Methods-Ansatzes
- Projektziel 2: Durchführung einer querschnittlichen, quantitativen Befragung, in deren Rahmen individuelle Bedingungen der Akzeptanz bei Versicherten (zweistufige Einwohnermelderegisterstichprobe) und ÄrztInnen (Auswahlstichprobe) anhand beispielhafter, auf Basis des Kategorisierungsinstrumentes entwickelter KI-Anwendungsszenarien evaluiert wurden.
- Projektziel 3: Translation der Ergebnisse aus der Befragung und Ableitung praktischer Handlungsempfehlungen für den nutzergerechten Einsatz von KI-Systemen in der Versorgung im Rahmen eines Expertenworkshops.

Für Projektziel 1 konnte mittels einer umfangreichen Literaturanalyse und semistrukturierten Interviews mit Experten ($n = 7$) ein mehrdimensionales Kategorisierungsinstrument erarbeitet werden, aus dem drei beispielhafte KI-Anwendungsszenarien aus den Bereichen Prävention, Diagnostik und Therapie für die quantitative Befragung als Stimuli umgesetzt werden konnten. Für Projektziel 2 konnte im Rahmen der Akzeptanzstudie mit 1,259 Teilnehmern (157 ÄrztInnen, 1,102 Versicherte) gezeigt werden, dass die KI-Akzeptanz im Gesundheitswesen von mehreren Faktoren beeinflusst wird. Menschen bevorzugten menschliche Entscheidungen gegenüber KI-getroffenen, wobei die Akzeptanz je nach Anwendungsbereich (z.B. Insulinpumpe am höchsten) variierte. Versicherte zeigten generell eine höhere Akzeptanz als ÄrztInnen. Eine neu entwickelte Vier-Faktor-Skala erwies sich als valides Instrument zur Messung der individuellen KI-Akzeptanz. Wichtige Prädiktoren zum Verständnis interindividueller Unterschiede in der Ausprägung der Akzeptanz gegenüber den KI-Stimuli waren unter anderem die entwickelte Akzeptanzskala, die Technikmotivation und das Lebensalter, wobei sich die Einflussfaktoren zwischen ÄrztInnen und Versicherten unterschieden.

Im Rahmen des Expertenworkshops mit $n = 13$ Personen konnten unter Zuhilfenahme von Design-Thinking-Methoden Anforderungen in den Bereichen *Definition von KI*, *Rahmenbedingungen des KI-Einsatzes*, *Entwicklung von KI-basierten Lösungen* sowie *Implementierungen von KI-basierten Lösungen* identifiziert und priorisiert werden.

2. Beteiligte Projektpartner

Tabelle 1: Beteiligte Projektpartner und Verantwortlichkeiten

Institution	Personen	Verantwortlichkeit
Fraunhofer-Institut für Integrierte Schaltungen IIS	<ul style="list-style-type: none"> • Dr. Andreas Hamper (Projektleitung/Konsortialführung 01.08.21 – 30.11.21) • Martina Simon (Projektleitung/Konsortialführung 01.12.21 – 31.01.24) • Dr. Stephanie Schmitt-Rüth • PD Dr. Thomas Wittenberg • Dr. Bettina Williger • Prof. Dr. Alexander Martin 	<ul style="list-style-type: none"> • Projektleitung/Konsortialführung (AP 10) • Kategorisierung von KI-Anwendungen (AP 1) • Stimulusdesign (AP 2) • Rekrutierung und Durchführung der Studie mit ÄrztInnen (AP 6) • Auswertung der Studie mit ÄrztInnen (AP7) • Ableitung von Handlungsempfehlungen für die Praxis (AP 9)
Institut für Psychogerontologie (IPG) der Friedrich-Alexander-Universität Erlangen-Nürnberg	<ul style="list-style-type: none"> • Dr. Stefan Kamin (Verantwortung Methodik) • Dr. Anja Beyer • Prof. Dr. Frieder Lang (Projektleitung IPG) • Nora M. Degen • Helena Hoesch 	<ul style="list-style-type: none"> • Studiendesign (AP 3) • Rekrutierung und Durchführung der Studie mit Versicherten (AP 4) • Auswertung der Studie mit Versicherten (AP 5) • Synthese der Studienergebnisse (AP 8)
AOK Bayern	Assoziierter Partner	Begleitung und Beratung des Projektkonsortiums
Techniker Krankenkasse	Assoziierter Partner	Begleitung und Beratung des Projektkonsortiums
Sozialstiftung Bamberg	Assoziierter Partner	Begleitung und Beratung des Projektkonsortiums

Ansprechpartnerin für Rückfragen nach Projektende ist Martina Simon (martina.simon@iis.fraunhofer.de).

3. Projektziele

Die digitale Transformation und die damit verbundenen Möglichkeiten, klinische Daten in immer besserer Qualität und großer Menge zu erfassen, zu übertragen und zu verarbeiten, rückt zunehmend die Potentiale Künstlicher Intelligenz (KI) für die Versorgung im Rahmen der Gesetzlichen Krankenversicherung (GKV) in den Vordergrund. Der Einsatz von KI ermöglicht die Verarbeitung großer Datenmengen und kann dazu beitragen, komplexe Abläufe und Entscheidungsprozesse in der Gesundheitsversorgung zu optimieren (Auer, Hollenstein, & Reumann, 2019). Jedoch besteht auf Seiten von Patienten und ÄrztInnen gleichermaßen häufig große Unsicherheit, wofür KI im Gesundheitswesen eingesetzt werden kann, welche Risiken und Potentiale bestehen und wie die Qualität und Transparenz von KI-Anwendungen sichergestellt werden kann. Dabei kommt im Versorgungsalltag der Akzeptanz gegenüber der eingesetzten KI-Lösungen eine besondere Rolle zu. Nur wenn auf KI basierende Lösungen von den beteiligten Akteuren akzeptiert werden, können diese erfolgreich im Rahmen der

Versorgung implementiert werden (Stöber, Williger, Meerkamm, & Lang, 2012). Vor diesem Hintergrund ist es wichtig zu verstehen, welche Bedingungen die Akzeptanz von KI in der Gesundheitsversorgung beeinflussen.

Aus diesem Grund untersuchte das vorliegende Projekt die **Akzeptanz von KI-Anwendungen** in beispielhaften Versorgungsbereichen der GKV. Akzeptanz beschreibt hier die für die Nutzung technischer Innovationen erforderliche Einstellungsbildung (Davis, 1989), die sich beispielsweise durch den erwarteten Nutzen und die Alltagstauglichkeit von KI-Anwendungen manifestiert. Darüber hinaus erfordert die Analyse der KI-Akzeptanz eine dezidierte Betrachtung ethischer und sicherheitsrelevanter Aspekte wie Vertrauenswürdigkeit, Selbstbestimmtheit oder Datensicherheit im Kontext der Nutzung (Bundesregierung, 2018). Eine weitere Forschungsfrage ist, inwiefern die Akzeptanz gegenüber KI-Lösungen durch **individuelle Einflussfaktoren** (z.B. Alter, Bildung oder Technikaffinität) erklärt werden kann (Kamin, Lang, & Beyer, 2017; Seifert, Kamin, & Lang, 2020), um Handlungsempfehlungen für den nutzergerechten Einsatz von KI in Abhängigkeit unterschiedlicher Zielgruppen und Anwendungsbereiche zu identifizieren (Kamin et al., 2014; Kett, Ringler, Kamin, Lang, & Wartzack, 2015).

Das konzeptuelle und wissenschaftliche Vorgehen richtete sich an drei Projektzielen aus): **Projektziel 1** beinhaltete die Erstellung eines mehrdimensionalen und interdisziplinären **Kategorisierungsinstruments für KI-Anwendungen** in der medizinischen Versorgung. **Projektziel 2** umfasste die Durchführung einer querschnittlichen **Akzeptanzstudie** (quantitative Befragung mit $n = 1.536$ Versicherten und $n = 475$ ÄrztInnen in Nordbayern), in deren Rahmen die anhand des KI-Kategorisierungsinstruments entwickelten Stimuli in Form animierter Kurzvideos durch Versicherte und ÄrztInnen hinsichtlich ihrer individuellen Akzeptanz bewertet wurden. Hierbei wurde angenommen, dass sowohl hinsichtlich des KI-Anwendungsszenarios als auch hinsichtlich des Grades der KI-Autonomie sowie des Detaillierungsgrades der KI-Funktionsweise interindividuelle Unterschiede in der KI-Akzeptanz in Abhängigkeit soziodemografischer Faktoren wie Alter, Geschlecht, Persönlichkeit, Gesundheitsverhalten, gesundheitlichen Ressourcen, Technikmotivation, Digitalkompetenz sowie räumlichem Kontext (z.B. Versorgungssituation, Gemeindegröße) messbar sind. **Projektziel 3** zielte auf die Translation der Ergebnisse aus der Akzeptanzstudie und die Ableitung praktischer **Handlungsempfehlungen** für den nutzergerechten Einsatz von KI-Systemen in der Versorgung ab.

4. Projektdurchführung

Die Durchführung des Projekts »KI-BA« war auf einen Zeitraum von 30 Monaten (01.08.2021-31.01.2024) ausgelegt und in 9 Arbeitspakete (AP) gegliedert. Abbildung 1 gibt einen Überblick über die Verknüpfung der einzelnen Arbeitspakete.

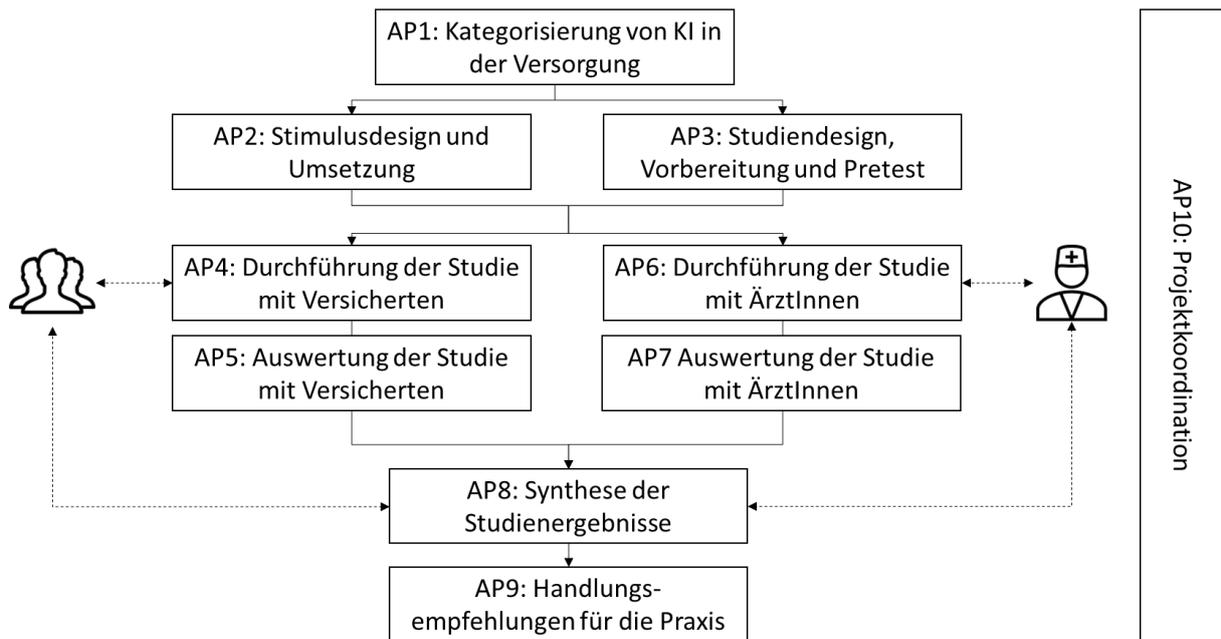


Abbildung 1: Schematischer Überblick über den Projektablauf

Die einzelnen Arbeitspakete wurden jeweils in enger Zusammenarbeit und Abstimmung zwischen dem Fraunhofer-Institut für Integrierte Schaltungen IIS und dem Institut für Psychogerontologie (IPG) der Friedrich-Alexander-Universität Erlangen-Nürnberg bearbeitet. Durch die umfangreiche Vorerfahrung des IPG hinsichtlich der Planung und Durchführung komplexer Studiendesigns, insbesondere auch im Hinblick auf Studien zur Nutzerpartizipation und Akzeptanz, war das IPG mit der Konzeption des Studiendesigns betraut, während das Fraunhofer IIS die Konsortialführung und damit die Projektkoordination übernahm. Die weiteren Arbeitspakete wurden ebenso jeweils federführend entweder vom IPG oder dem Fraunhofer IIS mit Zuarbeit des jeweils anderen Partners bearbeitet. Tabelle 2 stellt den zeitlichen Ablauf des Projekts basierend auf den einzelnen Arbeitspaketen sowie die Verantwortlichkeiten der beiden Konsortialpartner Fraunhofer IIS und IPG im Überblick dar.

5. Methodik

Das wissenschaftliche und methodische Vorgehen orientiert sich an den Projektzielen und ist in Abbildung 2 dargestellt. Grundanliegen des Projekts »KI-BA« war das Verständnis, wie die Akzeptanz gegenüber KI-Systemen bei Versicherten und ÄrztInnen ausgeprägt ist und welche Einflussfaktoren die individuelle Akzeptanz beeinflussen. In einem ersten Schritt wurden hierzu auf Grundlage des entwickelten **KI-Kategorisierungsinstruments** animierte Video-Sequenzen für KI-Anwendungen erstellt (Projektziel 1: Kategorisierung). Diese Stimuli wurden im Rahmen einer **querschnittlichen Akzeptanzstudie** durch Versicherte und ÄrztInnen bewertet (Projektziel 2: Akzeptanzstudie). Die letzte Phase des Vorgehens umfasste die Translation der generierten Erkenntnisse in die Praxis, indem **Handlungsempfehlungen** mit Experten erarbeitet werden (Projektziel 3: Handlungsempfehlungen).

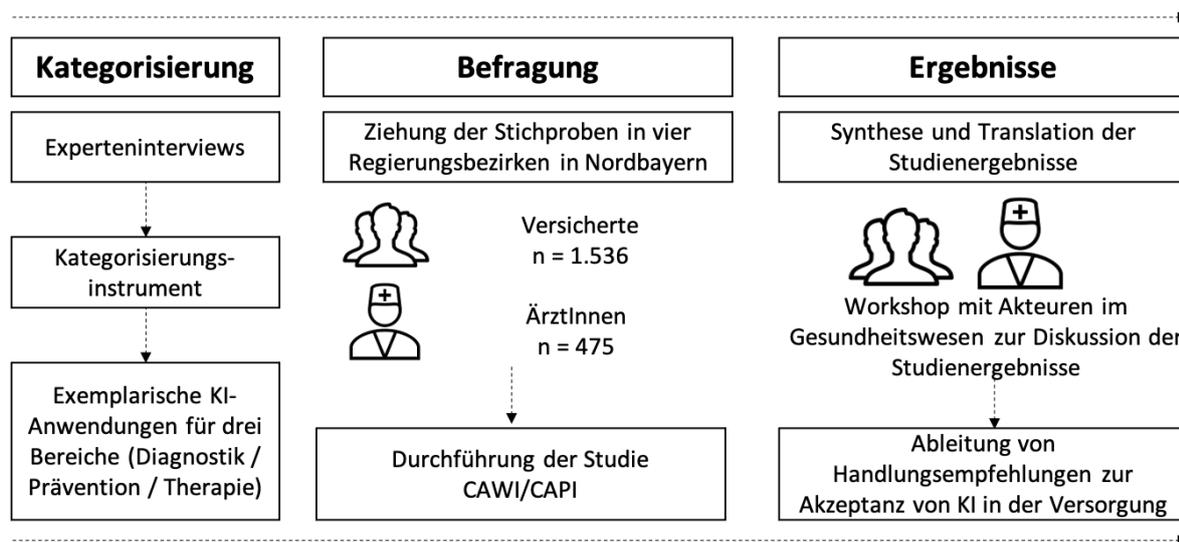


Abbildung 2: Übersicht über das wissenschaftliche und methodische Vorgehen

Für die einzelnen Projektziele wurde das in den nachfolgenden Kapiteln dargestellte methodische Vorgehen gewählt.

5.1 Methodik Kategorisierung

KI-Anwendungen für den Gesundheitsbereich sind vielfältig. Je nach Betrachtungswinkel kann die Kategorisierung z.B. nach Fachdisziplin (z.B. Radiologie, Endoskopie, Dermatologie, Geriatrie, Allgemeinmedizin), Einsatz in der Versorgung (z.B. Prävention, Diagnose, Therapie, Nachsorge oder Pflege), KI-Forschungsfeld (z.B. Automatisches oder adaptives Lernen, Explainable AI, Semantik oder mathematische Optimierung) oder Zielgruppe (z.B. Facharzt, Allgemeinarzt, Fachkraft, Patient, Verwaltung, Angehörige) unterschieden werden. Um diese vielfältigen Kategorisierungsmöglichkeiten von KI-Anwendungen für den Gesundheitsbereich angemessen berücksichtigen zu können, wurde ein vierstufiges methodisches Vorgehen gewählt (siehe Anhang 1): Die erste Stufe bestand aus einer **systematischen Literaturanalyse** unter Verwendung einschlägiger Datenbanken aus dem medizinischen, psychologischen und technischen Bereich, um einen Überblick über Kategorisierungsdimensionen der KI in der Versorgungslandschaft zu gewinnen und einen ersten Entwurf für ein Kategorisierungsinstrument zu erstellen. Hierzu wurden zunächst geeignete Schlüsselwörter identifiziert, Datenbanksuchen durchgeführt und Boolesche Operatoren zur Verfeinerung der Suchergebnisse angewendet. Der endgültige Suchausdruck, der in der Stichwortsuche verwendet wurde, um die Kategorien und Merkmale abzuleiten, lautete: title: („künstliche Intelligenz“ ODER „KI“ ODER „maschinelles Lernen“ ODER „ML“ ODER „tiefes Lernen“ ODER „DL“ ODER „neuronales Netzwerk“ ODER „NN“) UND („medizin*“ ODER „klinisch*“ ODER

„gesund*“ ODER „Pflege“)). Insgesamt wurden drei Online-Literaturdatenbanken durchsucht, die aus den Bereichen Medizin, Psychologie und Technologie ausgewählt wurden: Web of Science, Scopus und PubMed. Die Suche beschränkte sich nicht nur auf wissenschaftliche Literatur, sondern umfasste auch graue Literatur, um ein breites Spektrum an Wissen zu erfassen. Der aufeinanderfolgende Auswahlprozess umfasste das Screening und die Bewertung der Literatur basierend auf Relevanz und Qualität sowie das Entfernen von Duplikaten (siehe Abbildung 3). Mittels Konzeptmatrizen wurde die Analyse strukturiert und ein systematischer Vergleich der Literatur ermöglicht (siehe Anhang 2).

Im Rahmen der systematischen Literaturanalyse wurden die gefundenen Studien durch zwei unabhängige Bewerter evaluiert. Beide Bewerter folgten dabei in einem ersten Schritt den festgelegten Ein- bzw. Ausschlusskriterien, um eine konsistente und transparente Vorgehensweise zu gewährleisten (siehe Tabelle 3):

Tabelle 3: Ein- und Ausschlusskriterien für die Literatursuche

Einschlusskriterium	Beschreibung
Kategorisierungsansätze von KI in der Versorgung	Die Publikation fokussiert mögliche Kategorisierungsansätze von KI in der Versorgung und beschreibt Ausprägungen mind. einer Kategorie (<i>Input, Versorgungssäule, Fachgebiete, Nutzer, Technologie, Output</i>)
Ausschlusskriterium	Beschreibung
Ungültige Sprache	Die Publikation liegt in einer anderen Sprache als Englisch oder Deutsch vor.
Ungültige Studienart	Die Publikation stellt die Ergebnisse von klinischen Versuchen dar (z.B. trial, assessment, performance) und verfehlt damit das Ziel überblicksartiger Literaturquellen.
Nicht zielführender Fokus	Die Publikation bildet z.B. keine interdisziplinären Kategorien und Ausprägungen der KI in der Versorgung ab und stellt nur einseitige Bereiche und Akteure dar. Weiterhin werden komplexe KI-Techniken (Algorithmen, Entscheidungsbäume, Bayessche Netze etc.), spezifische Krankheitsphänomene (z.B. COVID) und Fachgebiete (Veterinärmedizin, Psychologie, Notarztwesen etc.) zu stark fokussiert, was im Sinne einer überblicksartigen Zielstellung zum Ausschluss der Publikationen führt.

In einem zweiten Schritt wurden die als relevant erachteten Publikationen von den beiden Bewertern in Konzeptmatrizen erfasst, strukturiert und kategorisiert, um verschiedene Arbeiten zueinander in Beziehung zu setzen. Im Falle von Bewertungsdiskrepanzen wurden diese systematisch diskutiert, bis ein Konsens erreicht wurde.

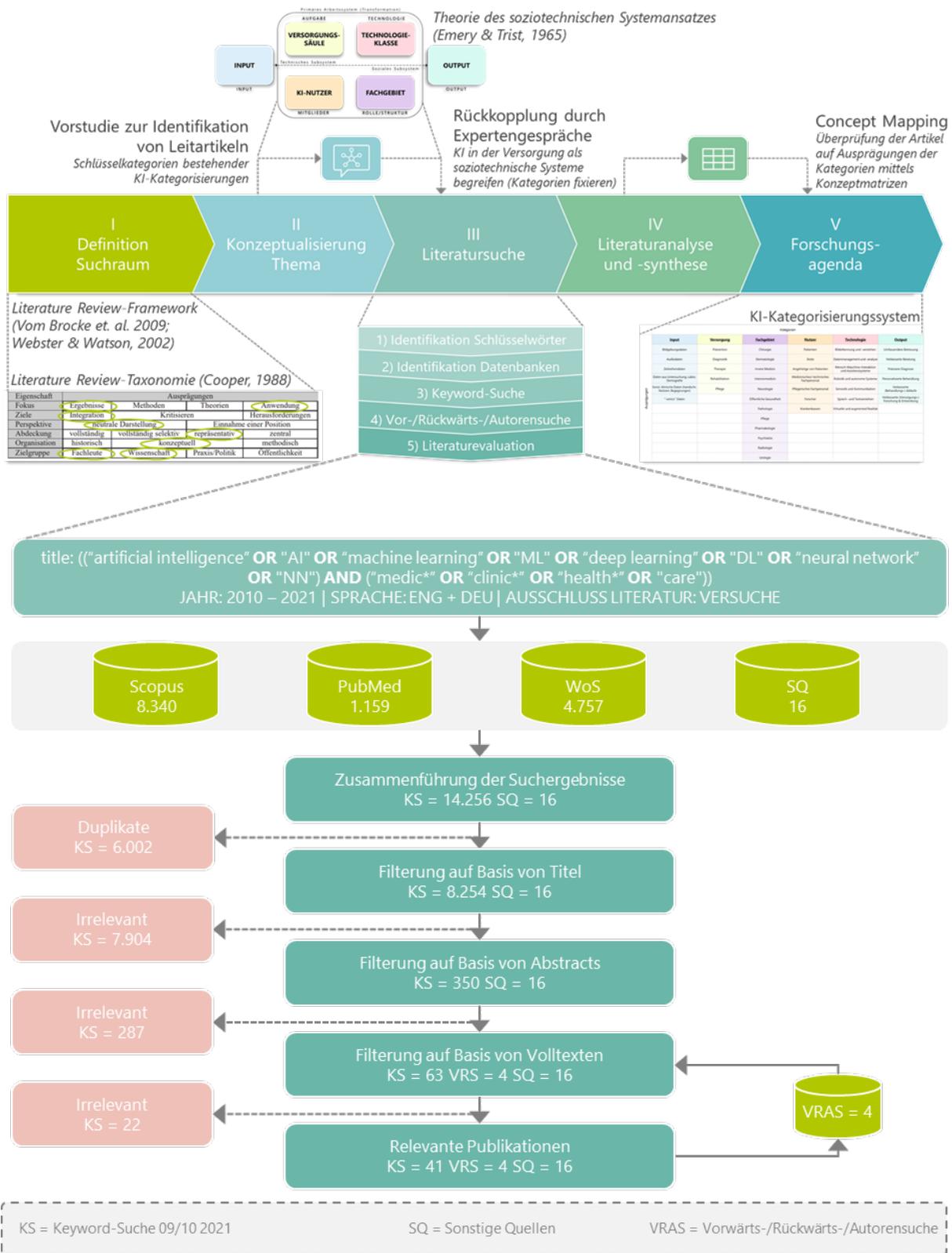


Abbildung 3: Systematische Literaturanalyse zur Erstellung eines ersten Kategorisierungsansatzes

Der aus der Literaturanalyse abgeleitete Kategorisierungsansatz wurde in der zweiten Stufe im Rahmen von **semistrukturierten Interviews** mit Experten aus der Gesundheitswirtschaft, Medizin, KI-Forschung, Medizintechnik und Versorgungsforschung diskutiert, erweitert und validiert (Zielgröße $n = 5$; Durchführungszeitraum 27.10.-07.12.2021).

Die Rekrutierung der Teilnehmer erfolgte aus dem Projektumfeld des Fraunhofer IIS und des IPG mittels direkter Ansprache und umfasste sowohl die Forschung zu Künstlicher Intelligenz als auch der Medizintechnik und Medizin. Da es sich bei den Experteninterviews um ein qualitatives Verfahren handelte, wurde auf eine Zufallsauswahl der Stichprobe verzichtet und stattdessen bei der Auswahl der Experten auf eine breite fachliche Streuung geachtet (Tausch & Menold, 2015). Das semistrukturierte, primär qualitativ orientierte Interview umfasste die nachfolgenden Bereiche:

- Demografika (Geschlecht, Alter, Beruf, spezifisches KI-Forschungsfeld, Beschäftigungsdauer mit KI, Einschätzung des eigenen KI-Expertenwissens auf einer 5er-Likert-Skala)
- Abgleich der Definition von KI des Probanden mit der im Projekt verwendeten Definition
- Freie Sammlung von KI-Versorgungsbeispielen inkl. Status (Entwicklung vs. Zulassung vorliegend), Vor- und Nachteilen sowie Chancen und Risiken des jeweiligen Beispiels.
- Vorstellung des auf Basis der Literaturanalyse entwickelten Kategorisierungsinstrumente und erste Eindrücke seitens der Probanden
- Kategorisierung eines selbst gewählten KI-Versorgungsbeispiels anhand des Kategorisierungsinstrumente inkl. Evaluation
- Versuch der Erklärung der eingangs genannten KI-Versorgungsbeispiele für einen KI-Neuling (z.B. Oma)
- Einschätzung der eigenen Nutzungswahrscheinlichkeit der KI-Versorgungsbeispiele auf einer 5er-Likert-Skala
- Evaluation des Kategorisierungssystems hinsichtlich der Aspekte Vollständigkeit, Verständlichkeit und Interdisziplinarität auf einer 5er-Likert-Skala

Im Rahmen eines Mixed-Methods-Ansatzes wurden die Experteninterviews in der dritten Stufe des methodischen Vorgehens um eine quantitative, als **Online-Fragebogen** angelegte Komponente ergänzt, um die im Rahmen der Experteninterviews generierten Beispiele für die nachfolgende Akzeptanzstudie mit Versicherten und ÄrztInnen zu klassifizieren und mittels einer Übersetzung in animierte Videosequenzen als Stimuli nutzbar zu machen. Der Online-Fragebogen (Durchführungszeitraum 02.-15.02.2022) adressierte dabei mit Blick auf das parallel entwickelte Studiendesign der Akzeptanzstudie die folgenden Bereiche bzw. Fragestellungen

- Bewertung der KI-Beispiele aus den Versorgungsbereichen Prävention, Diagnostik und Therapie hinsichtlich Verständlichkeit, Relevanz und Entlastungsfunktion auf einer 5er-Likert-Skala
- Auswahl und Priorisierung von je drei KI-Beispielen aus den Versorgungsbereichen Prävention, Diagnostik und Therapie, die sich nach eigener Einschätzung am besten dazu eignen, um sie gleichzeitig mit einer großen Anzahl an Versicherten und ÄrztInnen bezüglich KI-Akzeptanz zu testen
- Auswahl des Nutzungsszenarios bezogen auf die jeweils ausgewählten KI-Beispiele aus Arztsicht (keine Anwendung/Anwendung, wenn KI nur beobachtet oder fühlt, aber keine Entscheidung trifft (human in the loop)/Anwendung, wenn KI interpretiert oder evaluiert und Vorschläge macht (human on the loop)/Anwendung, wenn KI selbst entscheidet oder handelt (human out of the loop))
- Auswahl des Nutzungsszenarios bezogen auf die jeweils ausgewählten KI-Beispiele aus Patientensicht (keine Anwendung/Anwendung, wenn KI nur beobachtet oder fühlt, aber keine Entscheidung trifft (human in the loop)/Anwendung, wenn KI interpretiert

oder evaluiert und Vorschläge macht (human on the loop)/Anwendung, wenn KI selbst entscheidet oder handelt (human out of the loop))

- Nennung der gewünschten zusätzlichen Informationen, die aus Arzt- bzw. Patientensicht nötig wären, um der KI in den gewählten Beispielen die alleinige Entscheidung bzw. Handlung zu überlassen.

Die letzte Stufe des methodischen Vorgehens umfasste schließlich die **Erstellung animierter Videosequenzen** auf Basis der Ergebnisse aus dem Online-Fragebogen als Stimuli für die nachfolgende Akzeptanzstudie sowie den **Pretest** mit ÄrztInnen und Versicherten (Zielgröße $n = 10$; Durchführungszeitraum 08.-12.08.2022), um ein Verständnis der Stimuli sicherstellen und mögliche Verzerrungen durch das Design ausschließen zu können. Die Ziehung der Stichprobe für den Pretest folgte einem bewussten Auswahlverfahren im Rahmen einer direkten persönlichen Ansprache mit dem Ziel eines möglichst heterogenen Teilnehmerbilds. Der Pretest selbst erfolgte im Rahmen quantitativer und qualitativer Beurteilungen in einem Workshopformat, wobei verschiedene Methoden (Verhaltensbeobachtung, lautes Denken, Fragebogen) miteinander kombiniert wurden, um die Eignung des Materials zu bewerten. Die Auswertung der Pretests erfolgte basierend auf einer Clusterbildung, durch die sowohl Mehrfach- und Einfachnennungen als auch gegensätzliche Einschätzungen der Probanden identifiziert werden konnten. Die geclusterten Ergebnisse wurden im Anschluss von der Kommunikationsagentur Kalsperger Consulting, die im Rahmen eines Unterauftrags mit der Erstellung der Stimuli beauftragt wurde, in die finale Version der Stimuli eingearbeitet.

5.2 Methodik Akzeptanzstudie

Ziel der Akzeptanzstudie (Durchführungszeitraum 17.10.2022 – 28.02.2023) war es zu verstehen, wie die Akzeptanz gegenüber KI-Systemen bei Versicherten und ÄrztInnen ausgeprägt ist und welche Einflussfaktoren die individuelle Akzeptanz beeinflussen. Für die Beantwortung der Fragestellung war eine querschnittliche Akzeptanzstudie vorgesehen. Die Durchführung wiederholter Befragungen (Panelstudie) war nicht erforderlich, da die Analyse von Veränderungsprozessen über die Zeit nicht im Fokus des vorliegenden Projekts stand. Entsprechend entfallen im Folgenden auch Angaben zum Drop-Out. Die Studie wurde je nach technischen Möglichkeiten der Probanden als persönlich-mündliche Befragung (CAPI) oder als Online-Fragebogen (CAWI) durchgeführt. Durch diese Methodenkombination konnte die jeweilige zugrunde liegende Grundgesamtheit besser repräsentiert werden (z.B. bei fehlendem Internetzugang). Die folgenden Abschnitte geben einen Überblick über Stichprobendesign und Fallzahlberechnung je Zielgruppe sowie das generelle Studiendesign, Erhebungsinstrumente, Hypothesen sowie die Auswertungsmethodologie.

Stichprobendesign Versicherte

Die Ziehung der Versicherten-Stichprobe ($n = 1.536$ Personen über 18 Jahre in Privathaushalten) erfolgte durch eine mehrstufige Zufallsauswahl auf Basis lokaler Melderegister in Nordbayern. In der ersten Stufe wurden 50 Gemeinden zufällig und proportional zur Bevölkerungszahl ausgewählt. Dieses Vorgehen gewährleistet eine breite regionale Repräsentanz der Teilnehmenden und stellt zugleich eine akzeptable Balance zwischen Klumpungsfaktor und Kostenaufwand sicher. Die zweite Stufe umfasste die Ziehung einer disproportional nach Alter geschichteten Zufallsstichprobe. Dazu wurden aus den Auswahleinheiten (Gemeinden) der ersten Stufe gleich große Zufallsstichproben aus drei Altersgruppen gezogen (18-39 Jahre, 40-59 Jahre, 60 Jahre und älter). Dieses Vorgehen wurde gewählt, um extreme Zufallsstichproben (z.B. Überrepräsentation jüngerer Teilnehmer) zu vermeiden und sicherzustellen, dass Teilnehmende unterschiedlicher Altersgruppen vertreten waren. Um in die Studie eingeschlossen zu werden, mussten Personen über 18 Jahre alt sein. Nicht eingeschlossen wurden Personen, die (dauerhaft) in stationären Einrichtungen leben. Die gezogenen Personen wurden postalisch kontaktiert und zur Teilnahme an der Online-

Befragung eingeladen (CAWI) bzw. im Falle eines nicht vorhandenen Internetzugangs im Rahmen eines Interviews in der häuslichen Umgebung befragt (CAPI).

Fallzahlberechnung Versicherte

Für die Fallzahlberechnung wurden zwei zentrale Parameter zu Grunde gelegt: (1) ein angestrebtes Konfidenzniveau (Z) von 95% und (2) eine tolerierte Fehlerwahrscheinlichkeit (E) von 2.5%. Die Berechnung erfolgte nach der Formel $n = p(100-p)z^2/E^2$ und ergab eine anzustrebende Gesamtstichprobe von $N = 1.536$ Personen ($n = 512$ Personen je Altersgruppe). Die Berechnung der Teststärke erfolgte anhand des unten beschriebenen Versuchsplans in einem $2 \times 2 \times 3$ -faktoriellen Design. Für die Berechnung wurden 12 Gruppen sowie 10 Kovariaten als Berechnungsgrundlage angenommen. In diesem Rechenmodell ist ein Stichprobenumfang von $N = 1.536$ Personen ausreichend, um einen kleinen Effekt ($f = .10$) mit einer Wahrscheinlichkeit von 77% zu entdecken.

Stichprobendesign ÄrztInnen

Die Grundgesamt der ÄrztInnen-Stichprobe umfasste das ambulante Versorgungsangebot der hausärztlichen und fachärztlichen Versorgung sowie in Krankenhäusern und Kliniken angestellte ÄrztInnen in Nordbayern. Bei ÄrztInnen handelt es sich um eine Subpopulation, für die klassische Zufallsverfahren nur mit hohem Aufwand zu realisieren sind. Aus diesem Grund folgte die Ziehung dieser Stichprobe einem bewussten Auswahlverfahren durch direkte Ansprache unterschiedlicher Akteure, um eine möglichst heterogene Stichprobe im Hinblick auf die Fachdisziplin, das Versorgungsumfeld, die Region und die Berufserfahrung abzubilden. Der angestrebte Stichprobenumfang für die ÄrztInnen lag bei $n = 475$ Personen.

Fallzahlberechnung ÄrztInnen

Für die Fallzahlberechnung wurden zwei zentrale Parameter zu Grunde gelegt: (1) ein angestrebtes Konfidenzniveau (Z) von 95% und (2) eine tolerierte Fehlerwahrscheinlichkeit (E) von 4,5%. Aufgrund des geringeren avisierten Stichprobenumfangs fällt die Fehlerwahrscheinlichkeit im Vergleich zur Versichertenstichprobe höher aus. Allerdings liegt diese mit unter 5% in einem akzeptablen und üblichen Bereich. Die Berechnung erfolgte nach der Formel $n = p(100-p)z^2/E^2$ und ergab eine anzustrebende Gesamtstichprobe von $N = 475$ Personen. Die Berechnung der Teststärke erfolgte anhand des unten beschriebenen Versuchsplans in einem $2 \times 2 \times 3$ - faktoriellen Design. Für die Berechnung wurden 12 Gruppen sowie 10 Kovariaten als Berechnungsgrundlage angenommen. In diesem Rechenmodell ist ein Stichprobenumfang von $N = 475$ Personen ausreichend, um einen mittleren Effekt ($f = .25$) mit einer Wahrscheinlichkeit von 98% zu entdecken.

Auswertung

Die Auswertung dieser umfangreichen Studie mit ihrer Vielzahl an Variablen wurde jeweils an die spezifischen Fragestellungen und Ziele der einzelnen Analysen angepasst. Um einen Einblick in das breite Spektrum der angewandten Methoden zu geben, wird das Vorgehen anhand von drei exemplarischen Studien dargestellt:

Studie 1: Zur Untersuchung der Akzeptanz gegenüber KI-Anwendungsbeispielen wurde eine mehrfaktorielle ANOVA mit einem $2 \times 2 \times 3$ Design durchgeführt. Die gewählten Faktoren ergeben sich aus den Projektergebnissen zur Kategorisierung (siehe 6.1.) und der Projektskizze. Die Struktur lässt sich wie folgt darstellen:

Faktor 1: KI-Autonomie (2 Stufen)

- Mensch entscheidet
- KI entscheidet

Faktor 2: KI-Erklärbarkeit (2 Stufen)

- Funktionsweise erklärt

- Funktionsweise nicht erklärt

Faktor 3: Art der KI-Anwendung (3 Stufen)

- EKG-Analyse
- Insulin-Dosierung
- Koloskopie-Befundung

Die Gruppenzugehörigkeit (ÄrztInnen vs. Versicherte) stellt dabei eine quasi-experimentelle Variable dar, die als zusätzlicher Zwischen-Subjekt-Faktor in die Analyse einbezogen wird. Dies resultiert in 12 verschiedenen experimentellen Bedingungen ($2 \times 2 \times 3 = 12$). Ohne Berücksichtigung von weiteren Kovariaten, ermöglicht dieses Design die Analyse von:

- Drei Haupteffekten
- Drei Zweifach-Interaktionen
- Eine Dreifach-Interaktion

Mit diesem Design können sowohl die isolierten Effekte der einzelnen Faktoren als auch deren Wechselwirkungen auf die KI-Akzeptanz untersucht werden, wobei die natürliche Gruppierung der Teilnehmer berücksichtigt wird.

Studie 2: Die Entwicklung und Validierung des KI-Akzeptanz-Instruments erfolgte in einem mehrstufigen Prozess, der sich über eine Vor- und eine Hauptstudie erstreckte.

In der ersten Phase der Vorstudie wurde auf Basis einer umfassenden Literaturanalyse ein theoretisches Modell mit vier Dimensionen der KI-Akzeptanz entwickelt. Ein interdisziplinäres Forschungsteam generierte initial 45 Items, die sich gleichmäßig auf die vier postulierten Dimensionen verteilten. Die Items wurden mit standardisierten Einleitungssätzen formuliert. Die psychometrische Überprüfung erfolgte mittels Online-Befragung von Leistungsempfängern und -erbringern. Zur Itemselektion und Validierung wurde ein sequenzieller Analyseprozess durchgeführt: Zunächst wurde die Faktorenstruktur durch explorative Faktorenanalysen untersucht, wobei die Analysen simultan für beide Stichproben durchgeführt wurden. Die Itemselektion folgte dabei standardisierten psychometrischen Kriterien, um zu einer ökonomischen Kurzversion mit optimalen Messeigenschaften zu gelangen. In der Hauptstudie wurde das zuvor entwickelte Instrument einer umfassenden Validierung unterzogen. Die konfirmatorische Faktorenanalyse diente der Überprüfung der Faktorenstruktur und der Messinvarianz zwischen verschiedenen Gruppen. Dabei wurde sowohl die Vergleichbarkeit zwischen Versicherten und ÄrztInnen als auch zwischen verschiedenen Altersgruppen getestet. Zur Analyse der prädiktiven Validität wurde ein Strukturgleichungsmodell spezifiziert, das neben den vier Akzeptanzdimensionen auch relevante Kovariaten einbezog. Als zentrale abhängige Variable wurde der TAM-Gesamt erfasst, der die Akzeptanzbewertung der in Studie 1 gezeigten Stimuli misst. Zusätzlich wurden die wahrgenommene Technologiekompetenz sowie demografische Variablen als potenzielle Einflussfaktoren berücksichtigt. Diese systematische Validierungsstrategie ermöglichte die Entwicklung eines psychometrisch fundierten Instruments zur Erfassung der KI-Akzeptanz im Gesundheitswesen, das sowohl für Forschungszwecke als auch für die praktische Anwendung geeignet ist.

Studie 3: Die Analyse der Einflussfaktoren auf die KI-Akzeptanz (TAM-Gesamtwert) erfolgte durch einen methodisch differenzierten Ansatz des maschinellen Lernens. Zur Identifikation der relevantesten Prädiktoren wurden drei verschiedene regularisierte Regressionsverfahren eingesetzt:

- LASSO (Least Absolute Shrinkage and Selection Operator)
- Elastic Net
- Square-root LASSO

Diese Verfahren wurden speziell aufgrund ihrer Fähigkeit ausgewählt, mit Multikollinearität umzugehen und eine automatische Variablenselektion durchzuführen. Sie ermöglichen die Entwicklung sparsamer Modelle, indem sie irrelevante Prädiktoren durch Schrumpfung der Koeffizienten gegen Null eliminieren. Ergänzend wurde ein Random Forest als nicht-parametrisches Ensemble-Lernverfahren implementiert. Dieser Ansatz ermöglicht die Erfassung nicht-linearer Beziehungen und Interaktionseffekte zwischen den Prädiktoren, die von linearen Modellen möglicherweise nicht erkannt werden. Die Robustheit der Modelle wurde durch eine 5-fache Kreuzvalidierung sichergestellt. Dieses Validierungsverfahren ermöglicht eine Einschätzung der Modellgüte und minimiert das Risiko einer Überanpassung. Die Kombination verschiedener Analysemethoden mit systematischer Validierung gewährleistet dabei eine umfassende und reliable Identifikation der bedeutsamsten Einflussfaktoren auf den TAM-Gesamtwert.

Aufbau der Befragung

Der Aufbau der Befragung ist in Abbildung 3 dargestellt. Die erste Seite umfasste die Erhebung grundlegender soziodemographischer Informationen. Danach wurden die Probanden gebeten, sich ein Video zur Einführung in die KI-Thematik anzusehen. Nach diesem Video erfolgte die Erhebung der KI- Akzeptanzskala, die im Rahmen des Projektes entwickelt wurde, um interindividuelle Unterschiede in der Bewertung der KI-Lösungen zu erklären. Das Design folgte einem Between-Subject-Ansatz, bei dem die Teilnehmenden per Zufallsprinzip einer der experimentellen Bedingungen zugewiesen wurden. Die technische Implementation erfolgte über die SoSci Survey-Plattform, die eine automatisierte Randomisierung ermöglicht. Innerhalb jeder Experimentalbedingung wurde zusätzlich die Reihenfolge der zu evaluierenden Videosequenzen randomisiert. Die Evaluation erfolgte anhand klassischer Instrumente aus der Technikakzeptanz- sowie der UX-Forschung, die gleichzeitig als zentrale abhängige Variablen fungierten. Nach der Evaluation wurden die Probanden gebeten, weitere Fragen zu potenziellen Einflussfaktoren der KI-Akzeptanz zu beantworten. Folgende Instrumente wurden erhoben:

Table 4: Bei Versicherten und ÄrztInnen erfasste Konstrukte

Konstrukt	Quelle
Geburtsjahr	
Geschlecht	
Gemeindegröße	BBSR
Postleitzahl	
Technikmotivation	Kamin & Lang (2013)
KI-Akzeptanz in der Gesundheitsversorgung	Eigene Entwicklung
Allgemeine KI-Akzeptanz	Sindermann et al. (2021)
Technikakzeptanz	Davis et al. (1989)
Big Five	Lang et al. (2005)
Zeitperspektive	Rohr et al. (2017), Gellert et al. (2012)
Obsoleszenzerleben	Brandstädter & Wentura (1994)
Subjektives Alterserleben	Lawton (1975)
Novelty	Eigene Items
Neugier	Kashdan et al (2018)

Konstrukt	Quelle
Lebenszufriedenheit	Analog zu diversen Panelstudien
Subjektive Gesundheit	Analog zu SOEP
Einsamkeit	Hughes et al. (2004)
Ängstlichkeit und Depression	Löwe et al. (2010)
Vorsorge	DEAS
Zufriedenheit mit Beziehungen	Angelehnt an SOEP, NRW80+, DEAS, VW Studie
Soziale Unterstützung	Angelehnt an SOEP, NRW80+, DEAS
Anzahl wichtiger Beziehungen	Angelehnt an SOEP, NRW80+, DEAS

Tabelle 5: Nur bei Versicherten erfasste Konstrukte

Konstrukt	Quelle
Höchster Bildungsabschluss	SHARE, DEAS, SOEP
Erwerbstätigkeit	Kühne et al. (2020), SOEP-CORE
Haushaltsgröße	Lang et al. (2022)
Erreichbarkeit Versorgungsangebote	Bonaiuto et al. (1999)
COVID-19 Infektion	Analog zu diversen Paneldaten
Funktionale Beeinträchtigungen	Verbrugge (1997), Robine et al. (2003), Van Oyen et al. (2006)
Diagnosen	DEAS
Krankenhausaufenthalte	Analog zu diversen Panelstudien
Arztkontakte	DEAS
Krankenversicherung	Eigene Items
Arzt-Patienten-Beziehung (Patientensicht)	WFPTS Skala (Kurzversion)
Internetnutzung	DEAS

Tabelle 6: Nur bei ÄrztInnen erfasste Konstrukte

Konstrukt	Quelle
Facharzt	SHARE, DEAS, SOEP
Promotion	Eigene Items
Berufliche Position	Marburger Bund
Beschäftigungsverhältnis	Marburger Bund
Patientengruppe	Eigene Items
Arbeitsbedingungen	Marburger Bund
Arbeitsbelastung	Marburger Bund
Zeit für Patienten	Marburger Bund
Arzt-Patienten-Beziehung (ÄrztInnensicht)	Eigene Items

Konstrukt	Quelle
Berufswechsel	Marburger Bund
Digitalisierungsgrad in Praxis/Einrichtung	Marburger Bund
IT-Schulungen im Beruf	Marburger Bund
genutzte Technologien im Beruf	Bitkom Umfrage
KI-Erfahrung im Beruf	Eigene Items

Nach Präsentation der KI-Stimuli wurde ein Manipulationscheck durchgeführt. Die Teilnehmenden wurden gebeten, den medizinischen Anwendungskontext der gezeigten KI-Anwendung zu identifizieren. Hierzu wurde eine Single-Choice-Frage mit sieben Antwortoptionen verwendet, die neben den tatsächlich verwendeten Anwendungskontexten (Diabetes, Koloskopie, EKG) auch Distraktoren (Parkinson, Zahnwurzelbehandlung) sowie Ausweichoptionen ("Keiner der aufgezählten", "Das weiß ich nicht") enthielt. Dies ermöglichte die Kontrolle, ob die experimentelle Manipulation erfolgreich war.

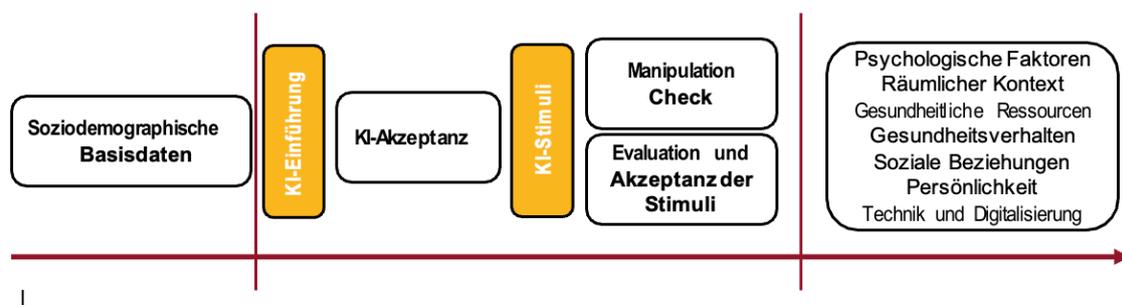


Abbildung 4: Aufbau der Befragung

Auswahl der Prädiktoren

Die Auswahl und Analyse von Prädiktoren unterscheidet sich methodisch bedingt zwischen den verschiedenen Studiendesigns.

Studie 1 (experimentelle Studien): Das faktorielle Design mit Varianzanalysen ermöglicht die systematische Untersuchung kausaler Zusammenhänge für ausgewählte Faktoren. Durch die kontrollierte Variation von KI-Autonomie, KI-Erklärbarkeit und Anwendungsart können spezifische Hypothesen über Wirkzusammenhänge getestet werden. Diese hohe interne Validität wird jedoch durch eine reduzierte ökologische Validität erkauft, da nur ein begrenzter Realitätsausschnitt abgebildet werden kann.

Studie 2 (Instrumentenentwicklung): Die psychometrische Entwicklung des KI-Akzeptanz-Instruments folgt einem konfirmatorischen Ansatz, der theoriegeleitet spezifische Konstruktdimensionen untersucht. Durch die systematische Reduktion auf zentrale Faktoren (Nützlichkeit, Kontrolle, Emotion, Vertrauen) wird ein fokussiertes Messinstrument geschaffen, das die Kernaspekte der KI-Akzeptanz reliabel erfasst.

Studie 3 (Machine Learning Ansätze): Die ML-Modelle verfolgen einen grundlegend anderen Ansatz, da sie nicht auf Hypothesenprüfung oder Populationsschlüsse abzielen, sondern auf die Identifikation von Vorhersagemustern. Dies ermöglicht die simultane Analyse einer Vielzahl potenzieller Einflussfaktoren aus verschiedenen Bereichen und somit den Einschluss nahezu aller in der Befragung erfassten Prädiktoren.

Diese drei Ansätze ergänzen sich gegenseitig: Die Instrumentenentwicklung schafft die psychometrische Grundlage für die valide Messung der KI-Akzeptanz. Die experimentellen Studien identifizieren kausale Mechanismen für zentrale Einflussfaktoren. Die ML-Modelle

liefern ein umfassendes Bild der komplexen Prädiktorstruktur unter Einbeziehung vielfältiger Kontextvariablen. Diese methodische Triangulation ermöglicht ein differenziertes Verständnis der KI-Akzeptanz. Während das entwickelte Instrument und die experimentellen Studien gezielt einzelne Aspekte untersuchen, können die ML-Analysen die Komplexität realer der empirischen Realität besser abbilden.

Einen detaillierten Überblick über den Aufbau der Befragungen und die erhobenen Instrumente bieten Anlage 1 und Anlage 2 sowie Anhang 4.

5.3 Methodik Handlungsempfehlungen

Ziel der letzten Projektphase war die Translation der Studienergebnisse in konkrete Handlungsempfehlungen für den Einsatz von KI in der Versorgung, um die Ergebnisse der Akzeptanzstudie auch für Akteure aus der Praxis nutzbar zu machen. Hierzu wurde ein spezifisches Workshopkonzept („Expertenworkshop“) erstellt, in dem mit KI-Themenexperten zum einen Implikationen für die nutzerzentrierte Entwicklung und Gestaltung von KI-Systemen, zum anderen Strategien zum Einsatz von KI-Systemen in der GKV erarbeitet und in konkrete Handlungsempfehlungen überführt werden. Ziel ist es, den identifizierten Akzeptanzrisiken und Barrieren bei der Einführung von KI-Technologien zu begegnen. Um diesen Workshopzielen gerecht werden zu können, wurde die Zusammensetzung einer möglichst heterogenen und repräsentativen Teilnehmergruppe von Experten, bestehend insb. aus ÄrztInnen, Patientenvertretern, Vertretern der GKV, Politik, Ethikexperten sowie Technikexperten, angestrebt (Zielgröße $n = 15$, Durchführung 20.07.23). Die Rekrutierung erfolgte dabei über eine Akquise im Rahmen der in der Kategorisierungsphase erfolgten Experteninterviews sowie über eine direkte persönliche Ansprache bei schwieriger zu erreichenden Wunschexperten, wie z.B. Ethikexperten, Vertreter aus der Politik, Technikexperten oder ÄrztInnen. Da es sich beim geplanten Expertenworkshop um ein qualitatives Verfahren handelte, war keine Zufallsauswahl der Stichprobe nötig, vielmehr wurde bei der Auswahl der Experten darauf geachtet, dass sie für die Workshopziele möglichst eine Vielfalt an fachlichem Input liefern können (Tausch & Menold, 2015). Der Workshop selbst wurde auf Basis eines Fokusgruppen-Konzepts durchgeführt und folgte einer methodischen Trias (Hackel & Klebl, 2008): In einem ersten Schritt wurden der gesamten Expertenrunde hierzu die Ergebnisse der Akzeptanzstudie vorgestellt. Im nächsten Schritt erarbeiteten die Experten in moderierten Kleingruppen mittels Design-Thinking-Methoden eine Sammlung konkret in die Praxis umsetzbarer Implikationen. Hierzu wurden den Teilnehmenden verschiedene Triggerbegriffe als Anhaltspunkt vorgelegt. Diese wurden in Anlehnung an Lucke (1995) gewählt. Berücksichtigt wurden dabei die drei Konzepte Akzeptanzobjekt, Akzeptanzsubjekt und Akzeptanzkontext. Diese wurden zum einfacheren Verständnis durch die Begriffe Technologie, Individuum und Nutzungskontext ersetzt. Im letzten Schritt folgte eine konsolidierende Diskussion in der Gesamtgruppe mit dem Ziel einer ersten groben Sammlung praktischer Anforderungen, die auf Post-its notiert wurden. Darüber hinaus wurde im Rahmen eines Kurzfragebogens Geschlecht, Alter, Beruf, Dauer der Beschäftigung und Technikbereitschaft (Neyer, Felber & Gebhard, 2016) erhoben.

6. Projektergebnisse

Für die einzelnen Projektziele können die in den nachfolgenden Kapiteln dargestellten Ergebnisse berichtet werden.

6.1 Projektergebnisse Kategorisierung

Die relevanten Ergebnisse aus der Literaturanalyse mit $n = 61$ Quellen (siehe Anhang 2) wurden in einen ersten Entwurf eines KI-Kategorisierungsinstrumentes überführt, das auf den

Kategorien *Input*, *Versorgungssäule*, *Fachgebiet*, *Nutzer*, *KI-Technologie* und *Output* basiert (siehe Abbildung 5).

Dieser Entwurf des KI-Kategorisierungsinstruments wurde mit $n = 7$ ExpertInnen aus der Gesundheitswirtschaft, Medizin, KI-Forschung, Medizintechnik und Versorgungsforschung im Rahmen der unter 5.1 dargestellten semistrukturierten Online-Interviews evaluiert (Dauer pro Interview ca. 60 Minuten).

		Kategorien					
		Input	Versorgung	Fachgebiet	Nutzer	Technologie	Output
Ausprägungen	Bildgebungsdaten	Prävention	Chirurgie	Patienten	Bildererkennung und -verstehen	Umfassendere Betreuung	
	Audiodaten	Diagnostik	Dermatologie	Ärzte	Datenmanagement und -analyse	Verbesserte Beratung	
	Zeitreihendaten	Therapie	Innere Medizin	Angehörige von Patienten	Mensch-Maschine-Interaktion und Assistenzsysteme	Präzisere Diagnose	
	Daten aus Untersuchung, Labor, Demografie	Rehabilitation	Intensivmedizin	Medizinisches/-technisches Fachpersonal	Robotik und autonome Systeme	Personalisierte Behandlung	
	Sonst. klinische Daten (handschr. Notizen, Begegnungen)	Pflege	Neurologie	Pflegerisches Fachpersonal	Sensorik und Kommunikation	Verbesserte (Behandlungs-) Abläufe	
	"-omics" Daten		Öffentliche Gesundheit	Forscher	Sprach- und Textverstehen	Verbesserte (Versorgungs-) Forschung & Entwicklung	
			Pathologie	Krankenkassen	Virtuelle und augmented Realität		
			Pflege				
			Pharmakologie				
			Psychiatrie				
		Radiologie					
		Urologie					

Abbildung 5: Erster Entwurf eines Kategorisierungssystems

Die auf Tonband bzw. Video aufgezeichneten Interviews wurden sowohl im Rahmen einer qualitativen Inhaltsanalyse in Anlehnung an Mayring (2015) als auch deskriptiv ausgewertet. So konnten einerseits die von den Experten im genannten exemplarischen KI-Vertreter anhand des Kategorisierungssystems systematisiert und andererseits das KI-Kategorisierungsinstrument hinsichtlich der Aspekte Vollständigkeit, Verständlichkeit und Interdisziplinarität auf einer 5-stufigen Likert-Skala (1= trifft überhaupt nicht zu, 5 = trifft voll zu) bewertet werden. Abbildung 6 zeigt die insgesamt 19 in den Interviews erarbeiteten KI-Beispiele und deren Einordnung in das Kategorisierungssystem.

Titel	Input	werden mittels	KI-basierte Technologie	durch Nutzer	der Fachgebiet	zur Versorgung	aufbereitet für	Output
Multimodale Zeitreihendaten (Fitnessshirt-Monitoring)	Zeitreihendaten	werden mittels	Biosignal-/Vitalwert-/Sensorikanalyse	durch Ärzte	der Inneren Medizin	zur Diagnostik	aufbereitet für	präzisere (Ausschluss-) Diagnosen.
KI-basierte Schmerzanalyse in der Pflege	Bilddaten, Zeitreihendaten sowie sonstige klinische Daten	werden mittels	Bildererkennung und -verstehen & Biosignal-/Vitalwert-/Sensorikanalyse	durch pflegerisches und durch medizinisches/-technisches Fachpersonal	der Intensivmedizin und Pflege	zur Diagnostik	aufbereitet für	verbesserte Detektionen von Anomalien und Auffälligkeiten, verbesserte Dokumentationen, präzisere (Ausschluss-) Diagnosen.
KI-basierte Detektion von Läsionen / Polypen	Bilddaten	werden mittels	Bildererkennung und -verstehen	durch Ärzte	der Inneren Medizin	zur Prävention	aufbereitet für	verbesserte Detektionen von Anomalien und Auffälligkeiten.
AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	Bilddaten, Audiodaten, Zeitreihendaten sowie sonstige klinische Daten	werden mittels	Bildererkennung und -verstehen, Biosignal-/Vitalwert-/Sensorikanalyse, Sprach- und Textverstehen sowie Virtual und Augmented Reality	durch pflegerisches Fachpersonal	der Pflege, Dermatologie sowie Inneren Medizin	zur Prävention, Therapie und Pflege	aufbereitet für	umfassendere Betreuungen, verbesserte, personalisierte Behandlungen, verbesserte (Behandlungs-) Abläufe, verbesserte Detektionen von Anomalien und Auffälligkeiten sowie verbesserte Dokumentationen.
Antibiotikagabe im Krankenhaus	Daten aus Untersuchungen, Laboren, Demografien	werden mittels	Biosignal-/Vitalwert-/Sensorikanalyse	durch Ärzte	der Intensivmedizin	zur Therapie	aufbereitet für	verbesserte (Behandlungs-) Abläufe
Genetischspezifizierte Antibiotikagabe im Krankenhaus	Anamnesedaten, Daten aus Untersuchungen, Laboren, Demografien sowie Omics-Daten	werden mittels	Biosignal-/Vitalwert-/Sensorikanalyse	durch Ärzte	der Intensivmedizin	zur Therapie	aufbereitet für	verbesserte (Behandlungs-) Abläufe
Ultraschallbasierte Flüssigkeitsgabe im Notarztesen	Bilddaten	werden mittels	Bildererkennung und -verstehen	durch Ärzte	des Notarztesens	Diagnostik zur und Therapie	aufbereitet für	präzisere (Ausschluss-) Diagnosen.
Frühzeitige Erkennung von Unterzuckerung auf der Intensivstation	Daten aus Untersuchungen, Laboren, Demografien	werden mittels	Mensch-Maschine-Interaktionen & Assistenzsysteme sowie Biosignal-/Vitalwert-/Sensorikanalyse	durch Ärzte und pflegerisches Fachpersonal	der Intensivmedizin	zur Therapie	aufbereitet für	präzisere (Ausschluss-) Diagnosen und verbesserte (Behandlungs-) Abläufe.
Frühzeitige Erkennung von Unterzuckerung für den Privatgebrauch	Daten aus Untersuchungen, Laboren, Demografien	werden mittels	Mensch-Maschine-Interaktionen & Assistenzsysteme sowie Biosignal-/Vitalwert-/Sensorikanalyse	durch Patienten und Angehörige	(der Intensivmedizin)	zur Therapie	aufbereitet für	präzisere (Ausschluss-) Diagnosen und verbesserte (Behandlungs-) Abläufe.
KI-basierte Erkennung von Brustkrebs	Bilddaten	werden mittels	Bildererkennung und -verstehen	durch Ärzte und medizinisches/-technisches Fachpersonal	der Radiologie	zur Prävention	aufbereitet für	präzisere (Ausschluss-) Diagnosen.
GKV-Routinedaten für Parkinson-Sturz-Prävention	Routinedaten der GKV	werden mittels	Big Data-Analyse	durch Personal der Krankenkassen	(der Neurologie)	zur Prävention	aufbereitet für	verbesserte Beratungen.
Longitudinale Datenauswertung / Krankheitsprädiktion	Bilddaten, Zeitreihendaten und Daten aus Untersuchungen, Laboren, Demografien	werden mittels	Bildererkennung und -verstehen, Biosignal-/Vitalwert-/Sensorikanalyse sowie Big Data-Analyse	durch Ärzte	der Inneren Medizin	zur Therapie	aufbereitet für	verbesserte, personalisierte Behandlungen und verbesserte (Behandlungs-) Abläufe.
Verletzungsdiaagnose paralympischer Athleten	Sonstige klinische Daten	werden mittels	Sprach- und Textverstehen	durch medizinisches/-technisches Fachpersonal	der Sportmedizin	zur Prävention	aufbereitet für	umfassendere Betreuungen und verbesserte Detektionen von Anomalien & Auffälligkeiten.
Autismustherapie	Bilddaten und Zeitreihendaten	werden mittels	Bildererkennung und -verstehen, Mensch-Maschine-Interaktion & Assistenzsysteme, Biosignal-/Vitalwert-/Sensorikanalyse sowie Big Data-Analyse	durch Fachpersonal und Angehörige von Patienten	der Psychologie	zur Therapie	aufbereitet für	verbesserte (Behandlungs-) Abläufe.
Blutbildanalyse Hämacam	Bilddaten	werden mittels	Bildererkennung und -verstehen	durch Ärzte und medizinisches/-technisches Fachpersonal	der Inneren Medizin	zur Diagnostik	aufbereitet für	präzisere (Ausschluss-) Diagnosen.
App für Kardiopatienten	Diskrete Vitaldaten, kontinuierliche Zeitreihendaten sowie Anamnesedaten	werden mittels	Biosignal-/Vitalwert-/Sensorikanalyse	durch Patienten und Ärzte	der Inneren Medizin	Diagnostik zur und Therapie	aufbereitet für	präzisere (Ausschluss-) Diagnosen, verbesserte, personalisierte Behandlungen, verbesserte (Behandlungs-) Abläufe und verbesserte Detektionen von Anomalien & Auffälligkeiten.
Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	Zeitreihendaten	werden mittels	Biosignal-/Vitalwert-/Sensorikanalyse	durch Ärzte	der Neurologie	Diagnostik zur und Therapie	aufbereitet für	verbesserte, personalisierte Behandlungen sowie verbesserte Detektionen von Anomalien & Auffälligkeiten.
Digitale Pathologie (medizinische Bildanalyse)	Bilddaten	werden mittels	Bildererkennung und -verstehen	durch Ärzte	der Pathologie	zur Diagnostik	aufbereitet für	präzisere (Ausschluss-) Diagnosen.
Leitlinienmodellierung / longitudinale Modellierung	Zeitreihendaten	werden mittels	Big Data-Analyse	durch Ärzte	der Onkologie	Diagnostik zur und Therapie	aufbereitet für	präzisere (Ausschluss-) Diagnosen und verbesserte, personalisierte Behandlungen.

Abbildung 6: Übersicht aller KI-Beispiele aus den Experteninterviews und Einordnung ins Kategorisierungssystem

Sowohl in Bezug auf Verständlichkeit als auch Abdeckung der wichtigsten Bereiche wurde das Kategorisierungssystem von den TeilnehmerInnen zusammenfassend im Mittel als sehr gut bewertet (siehe Tabelle 3). Aspekte, die insbesondere hinsichtlich der Vollständigkeit und Interdisziplinarität des Kategoriensystems zu einer lediglich durchschnittlichen Bewertung geführt haben, wurden als Änderungen für die finale Version des Kategorisierungssystems berücksichtigt, das in Abbildung 7 dargestellt ist.

Tabelle 7: Einschätzung des Kategorisierungssystems

	Mittelwert	Minimum/Maximum	N
Kategorisierungssystem verständlich	4,5 (SD=,54)	4/5	6
Kategorisierungssystem vollständig	4,0 (SD=,63)	3/5	6
Kategorisierungssystem deckt wichtigste Bereiche ab	4,83 (SD=,40)	4/5	6
Kategorisierungssystem wird interdisziplinärer Sichtweise gerecht	4,33 (SD=,81)	3/5	6

Anmerkung: Angegeben sind Mittelwerte inkl. des kleinsten bzw. größten Werts. SD entspricht der Standardabweichung des Mittelwertes. N gibt die Anzahl der Personen an, die die Frage beantwortet haben.

Kategorien					
Input	Versorgung	Fachgebiet	Nutzer	KI-basierte Technologie	Output
Bilddaten	Prävention	Chirurgie	Patienten	Bildererkennung und -verstehen	Umfassendere Betreuung
Audiodaten	Diagnose	Dermatologie	Ärzte	Big Data-Analyse	Verbesserte Beratung
Anamnesedaten	Therapie	Innere Medizin	Angehörige von Patienten	Mensch-Maschine-Interaktion und Assistenzsysteme	Präzisere (Ausschluss-) Diagnose
Routinedaten der GKV (Abrechnungsdaten)	Rehabilitation	Intensivmedizin	Medizinisches/-technisches Fachpersonal	Robotik und autonome Systeme	Verbesserte Detektion von Anomalien, Auffälligkeiten
Vitaldaten (diskrete Messung)	Pflege	Neurologie	Pflegerisches Fachpersonal	Biosignal- / Vitalwert- / Sensorikanalyse	Frühzeitige Intervention (Event-/Krankheitsvermeidung)
Zeitreihendaten (kontinuierliche Messung)		Neurologie	Forschende	Sprach- und Textverstehen	Verbesserte, personalisierte Behandlung
Daten aus Untersuchung, Labor, Demografie		Notarztwesen	Personal der Krankenkassen & Versicherungen	Virtuelle und augmented Realität	Verbesserte (Behandlungs-) Abläufe
Sonst. klinische Daten (handschr. Notizen, Begegnung)		Öffentliche Gesundheit	Personal der Krankenhausverwaltung	X-plainable AI	Verbesserte (Versorgungs-) Forschung & Entwicklung
"-omics" Daten		Pathologie	Personal sonstiger stationärer Einrichtungen		Verbesserte Dokumentation
Wissenschaftliche Datenlage (Fachliteratur, -studien, experten)		Pflege	Apotheker		
		Pharmakologie	Heilmittelerbringer (Physio-, Ergotherapeuten, Logopäden etc.)		
		Psychiatrie	Hilfsmittelverordner		
		Psychologie	Personal des Rettungsdienstes		
		Radiologie			
		Sportmedizin			
		Urologie			
		Onkologie			

Abbildung 7: Finales Kategorisierungssystem für KI im Gesundheitswesen

Um die in den Experteninterviews generierten KI-Anwendungsbeispiele (Abbildung 7) für die Akzeptanzstudie zu klassifizieren und nutzbar zu machen, wurden diese von $n = 5$ TeilnehmerInnen der Experteninterviews im Rahmen des Online-Fragebogens hinsichtlich ihrer Verständlichkeit, Relevanz und Entlastungsmöglichkeit sowie hinsichtlich ihrer Eignung für die Evaluation im Rahmen der Akzeptanzstudie sowohl mit der Zielgruppe der Versicherten als auch mit der Zielgruppe der ÄrztInnen bewertet. Im Rahmen einer quantitativen Auswertung auf Basis eines Mittelwertsvergleichs sowie qualitativer Aussagen konnten folgende KI-Beispiele als am geeignetsten für die Akzeptanzstudie identifiziert werden:

- a) Prävention: Bildbasiertes Erkennen von Darmpolypen
- b) Diagnostik: Sensorbasiertes Erkennen von Herz-Kreislauf-Erkrankungen
- c) Therapie: Sensorbasiertes Erkennen von Über- bzw. Unterzuckerung

Aufgrund der kleinen Stichprobe von wurde die Relevanz der Beispiele für die Gesundheitsversorgung wurde im Folgenden nochmals durch eine Literaturanalyse verifiziert.

Für eine detaillierte Darstellung der Arbeitsschritte und Ergebnisse hinsichtlich der Erstellung des Kategorisierungsinstruments und der Auswahl beispielhafter KI-Beispiele wird auf Anhang 1 verwiesen.

Die gewählten KI-Beispiele wurden im weiteren Verlauf im Rahmen eines Unterauftrags von der Kommunikationsagentur Kalsperger Consulting in animierte Video-Stimuli übersetzt. Für diesen Schritt wurde in enger Abstimmung mit den Konsortialpartnern ein Skript für die ausgewählten KI-Beispiele erstellt, das das in Kapitel 5.2 dargestellte Studiendesign der Akzeptanzstudie aufgreift: Pro KI-Beispiel aus den Bereichen Prävention (bildbasiertes Erkennen von Darmpolypen), Diagnostik (sensorbasiertes Erkennen von Herz-Kreislauf-Erkrankungen) und Therapie (sensorbasiertes Erkennen von Über- bzw. Unterzuckerung) wurden jeweils 4 Stimulusvarianten mit einer Länge von maximal 90 Sekunden erstellt, um UV 1 (AI Explainability high vs. low) und UV 2 (AI completeness high vs. low) abbilden zu können. Tabelle 4 gibt eine Übersicht über die erstellten Stimulusvideos.

Tabelle 8: Übersicht über Stimulusvideos

KI-Beispiel	UV 1 Explainability high (detaillierte Informationen über die Funktionsweise der KI)	UV 1 Explainability low (geringe Informationen über die Funktionsweise der KI)	UV 2 AI completeness high (KI entscheidet allein)	UV 2 AI completeness low (KI macht Vorschläge und Mensch entscheidet)
Prävention: Bildbasiertes Erkennen von Darmpolypen	Darm_KI-BA_KI- L.mp4 Darm_K-BA_ME- L.mp4	Darm_KI-BA_KI- K.mp4 Darm_K-BA_ME- K.mp4	Darm_KI-BA_KI- L.mp4 Darm_KI-BA_KI- K.mp4	Darm_KI-BA_ME- L.mp4 Darm_KI-BA_ME- K.mp4
Diagnostik: Sensorbasiertes Erkennen von Herz-Kreislauf- Erkrankungen	EKG_KI-BA_KI- L.mp4 EKG_K-BA_ME- L.mp4	EKG_KI-BA_KI- K.mp4 EKG_K-BA_ME- K.mp4	EKG_KI-BA_KI- L.mp4 EKG_KI-BA_KI- K.mp4	EKG_KI-BA_ME- L.mp4 EKG_KI-BA_ME- K.mp4
Therapie: Sensorbasiertes Erkennen von Über- bzw. Unterzuckerung	Insulin_KI-BA_KI- L.mp4 Insulin_K-BA_ME- L.mp4	Insulin_KI-BA_KI- K.mp4 Insulin_K-BA_ME- K.mp4	Insulin_KI-BA_KI- L.mp4 Insulin_KI-BA_KI- K.mp4	Insulin_KI-BA_ME- L.mp4 Insulin_KI-BA_ME- K.mp4

Neben den 12 Stimulusvideos wurde ein weiteres Video in Form einer kurzen Einführung in die KI-Thematik erstellt, das als Stimulus für die Erhebung der KI-Akzeptanzskala gemäß Studiendesign (siehe Kapitel 5.2) verwendet wurde.

Sämtliche erstellten Videos durchliefen im Zeitraum vom 08.-12.08.2022 einen Pretest mit $n = 10$ Versicherten und ÄrztInnen. Als Ergebnis des Pretests zeigten sich zielgruppenübergreifend teilweise deutliche Schwächen in den Stimuli hinsichtlich der Verständlichkeit und grafischen Darstellung, was in einer umfassenden Überarbeitung resultierte. Darüber hinaus wurde als Ergebnis aus den Pretest-Rückmeldungen pro Video zusätzlich eine vertonte Fassung erstellt, um Probanden die Auswahl zwischen einer unvertonten Fassung mit Untertiteln und einer vertonten Fassung mit Untertiteln anzubieten. Abbildung 8 zeigt vergleichend Ausschnitte aus den videobasierten Stimuli.



Abbildung 8: Ausschnitt aus dem Videostimulus "Bildbasiertes Erkennen von Darmpolypen" - links Pretest, rechts finale Version

Für eine Übersicht über die den Stimulusvideos zugrundeliegenden Textbausteine wird auf Anhang 3 verwiesen.

Alle Stimulusvideos stehen ferner unter folgendem Link zur Ansicht und zum Download zur Verfügung: [KI-BA Stimuli](#)

6.2 Projektergebnisse Akzeptanzstudie

Im Folgenden werden drei repräsentative Befunde zu unterschiedlichen Aspekten der Akzeptanzstudie dargestellt, die jeweils für wissenschaftliche Veröffentlichungen vorbereitet werden:

1. Studie zur Akzeptanz gegenüber KI-Anwendungsbeispielen (Video-Stimuli): Diese Analyse untersucht, wie verschiedene Faktoren wie KI-Autonomie, Erklärbarkeit und Art der Anwendung die Akzeptanz von exemplarischen KI-Anwendungen im Gesundheitswesen beeinflussen.
2. Validierung der KI-Akzeptanz-Skala: Hier wird die Entwicklung und Überprüfung eines neuen Messinstruments zur Erfassung der individuellen KI-Akzeptanz vorgestellt, einschließlich seiner Struktur und Anwendbarkeit bei verschiedenen Gruppen.
3. Analyse der Prädiktoren der KI-Akzeptanz: Diese Studie identifiziert mittels maschineller Lernverfahren die wichtigsten Einflussfaktoren auf die KI-Akzeptanz bei ÄrztInnen und Versicherten.

Die Ausgangsstichprobe für die dargestellten Studien umfasst insgesamt 157 ÄrztInnen (davon n = 0 CAPI) und 1.102 Versicherte (davon n = 5 CAPI). Die effektiven Stichprobengrößen in den einzelnen Analysen variieren jedoch in Abhängigkeit von fehlenden Werten auf den jeweils relevanten Variablen sowie den gewählten Verfahren zum Umgang mit diesen fehlenden Werten.

Die Rekrutierung der ÄrztInnen erfolgte über verschiedene Zugangswege, primär über Kliniken, Berufsverbände, Ärztenetze, persönliche Kontakte, Presse- und Social Media-Aufrufe sowie über den assoziierten Partner AOK Bayern, mit Schwerpunkt im Raum Nordbayern. Aufgrund der dezentralen Ansprache über insgesamt 236 Stellen bzw. Medien kann keine präzise Aussage über die Gesamtzahl der kontaktierten ÄrztInnen bzw. die Rücklaufquote getroffen werden.

Die Stichprobe der Versicherten wurde mittels mehrstufiger Zufallsauswahl auf Basis lokaler Melderegister in Nordbayern gezogen. In zwei Rekrutierungswellen wurden insgesamt 14.960 Einladungen zur Studie verschickt, was bei 1.102 Teilnehmenden einer Rücklaufquote von 7,37 % entspricht. Die variierenden Stichprobengrößen in den einzelnen Analysen werden in den jeweiligen Ergebnisdarstellungen gesondert ausgewiesen.

Die drei vorgestellten Teilstudien bilden ein kohärentes Forschungsprogramm zur Untersuchung der KI-Akzeptanz im Gesundheitswesen. Studie 1 untersucht mittels randomisierter KI-Anwendungsszenarien in Form animierter Kurzvideos die Akzeptanz gegenüber spezifischen KI-Anwendungen, wobei Faktoren wie KI-Erklärbarkeit, KI-Autonomie und verschiedene Use-Cases (Diabetes, EKG, Koloskopie) berücksichtigt werden. Diese Erkenntnisse fließen in Studie 2 ein, welche ein validiertes Messinstrument zur Erfassung der individuellen KI-Akzeptanz entwickelt. Die in Studie 2 erarbeitete KI-Akzeptanz-Skala dient wiederum als Grundlage für Studie 3, in der mittels maschineller Lernverfahren die wichtigsten Prädiktoren der KI-Akzeptanz identifiziert werden. Dabei werden auch die in Studie 1 identifizierten Einflussfaktoren sowie weitere potenzielle Einflussfaktoren wie räumlicher Kontext, soziale Beziehungen, gesundheitliche Ressourcen, Gesundheitsverhalten, Persönlichkeit und Technikaffinität berücksichtigt. Dieser Ansatz ermöglicht ein umfassendes Verständnis der KI-Akzeptanz im Gesundheitswesen, von konkreten Anwendungsszenarien über die Messung bis hin zur Vorhersage der Akzeptanz bei ÄrztInnen und Versicherten.

6.2.1 Studie 1: Akzeptanz gegenüber KI-Anwendungsbeispielen

Um die Akzeptanz von KI-Anwendungen zu untersuchen, wurde eine mehrfaktorielle ANOVA durchgeführt. Diese Methode ermöglicht es, den Einfluss mehrerer unabhängiger Variablen (Faktoren) auf eine abhängige Variable zu analysieren und potenzielle Interaktionseffekte zwischen den Faktoren zu identifizieren. In dieser Studie wurde eine 2x2x3 ANOVA mit folgenden Faktoren durchgeführt:

1. KI-Autonomie (2 Stufen: Mensch trifft Entscheidung vs. KI trifft Entscheidung)
2. KI-Erklärbarkeit (2 Stufen: lange vs. kurze Erklärung)
3. Art der KI-Anwendung (3 Stufen: EKG, Insulinpumpe, Koloskopie)

Die Gruppenzugehörigkeit (ÄrztInnen vs. Versicherte) stellt dabei eine quasi-experimentelle Variable dar, die als zusätzlicher Zwischen-Subjekt-Faktor in die Analyse einbezogen wurde.

Die abhängige Variable war der TAM-Gesamtwert, der die Akzeptanz der KI-Anwendungen misst. Das Hauptziel der ANOVA-Analyse war es, statistisch signifikante Unterschiede in der Akzeptanz von KI-Anwendungen zwischen den verschiedenen experimentellen Bedingungen und Gruppen zu identifizieren. Spezifisch sollte untersucht werden:

1. Ob die Faktoren KI-Autonomie, KI-Erklärbarkeit und Art der KI-Anwendung einen signifikanten Einfluss auf die Akzeptanz haben (Haupteffekte).
2. Ob es Unterschiede in der Akzeptanz zwischen ÄrztInnen und Versicherten gibt (Haupteffekt der Gruppenzugehörigkeit).
3. Ob Wechselwirkungen zwischen den Faktoren bestehen, die die Akzeptanz beeinflussen (Interaktionseffekte).

Um für mögliche Verzerrungen zu kontrollieren, wurden zusätzlich die Variablen Alter, Geschlecht und ATAI als Kovariaten in das Modell aufgenommen.

Stichprobe und Variablen

Die Stichprobe für diesen Teil der Akzeptanzstudie umfasste insgesamt 1,113 Teilnehmer, davon 122 ÄrztInnen und 991 Versicherte. Personen mit fehlenden Werten bei den einbezogenen Variablen wurden von den Analysen ausgeschlossen, um Verzerrungen in den Ergebnissen zu vermeiden. Das Durchschnittsalter der Gesamtstichprobe betrug 51 Jahre (SD = 16), wobei sich kein signifikanter Altersunterschied zwischen ÄrztInnen (M = 52, SD = 12) und Versicherten (M = 50, SD = 17) zeigte, $t(2, 363.38) = 1.01$, $p = .365$. Die Geschlechterverteilung in der Gesamtstichprobe war nahezu ausgeglichen mit 49% weiblichen und 51% männlichen Teilnehmern. Allerdings unterschied sich die Geschlechterverteilung signifikant zwischen den Gruppen, $\chi^2(2) = 8.36$, $p = .0153$, mit einem höheren Anteil männlicher Teilnehmer unter den ÄrztInnen (63%) im Vergleich zu den Versicherten (49%).

Die generelle Einstellung zu Künstlicher Intelligenz (KI) wurde mit dem ATAI (Attitude Towards Artificial Intelligence) erfasst, einer Skala von 0 (starke Ablehnung) bis 10 (starke Zustimmung). Der durchschnittliche ATAI-Wert betrug 6.6 (SD = 2.0) für die Gesamtstichprobe, ohne signifikanten Unterschied zwischen ÄrztInnen (M = 7.0, SD = 2.0) und Versicherten (M = 6.6, SD = 2.0), $t(2, 2223) = 2.41$, $p = .09$. Höhere Werte repräsentieren eine positivere generelle Einstellung gegenüber KI.

Die Technikakzeptanz wurde durch den TAM-Gesamtwert gemessen, der die Akzeptanz gegenüber den bewerteten KI-Anwendungen abbildet. Der durchschnittliche TAM-Gesamtwert für die Gesamtstichprobe betrug 3.2 (SD = 0.60), wobei ein signifikanter Unterschied zwischen ÄrztInnen (M = 3.1, SD = 0.64) und Versicherten (M = 3.2, SD = 0.59) festgestellt wurde, $t(2, 2223) = 3.18$, $p = .0419$.

Die experimentellen Bedingungen umfassten insgesamt 12 Video-Stimuli, die für die Analyse in drei Faktoren aufgeteilt wurden:

1. KI-Autonomie: "Mensch trifft Entscheidung" (50%) vs. "KI trifft Entscheidung" (50%)
2. KI-Erklärbarkeit: Video mit "langer Erklärung" (50%) vs. Video mit "kurzer Erklärung" (50%)
3. Anwendung: Video zeigt KI-Anwendung "EKG" (33%), "Insulinpumpe" (34%) oder "Koloskopie" (34%)

Es wurden keine signifikanten Unterschiede in der Verteilung dieser experimentellen Bedingungen zwischen ÄrztInnen und Versicherten gefunden (Autonomie: $\chi^2(2) = 0.57$, $p = .751$; Erklärbarkeit: $\chi^2(2) = 1.67$, $p = .435$; Anwendung: $\chi^2(4) = 3.07$, $p = .547$). Diese Faktoren ermöglichten eine differenzierte Analyse der Akzeptanz verschiedener KI-Anwendungsszenarien im Gesundheitsbereich. Eine detaillierte Beschreibung der Stichprobe und der Variablen ist in Tabelle 5 dargestellt.

Tabelle 9: Beschreibung der Stichprobe

Variablen	ÄrztInnen (N = 122)	Versicherte (N = 991)	Gesamt (N = 1,113)	t / χ^2	df	p
TAM-Gesamtwert *	3.1 (\pm 0.64)	3.2 (\pm 0.59)	3.2 (\pm 0.60)	3.18	2, 2223	.0419
Autonomie				0.57	2	.751
Mensch	65 (53 %)	492 (50 %)	557 (50 %)			
KI	57 (47 %)	499 (50 %)	556 (50 %)			
Erklärbarkeit				1.67	2	.435
Lang	54 (44 %)	500 (50 %)	554 (50 %)			
Kurz	68 (56 %)	491 (50 %)	559 (50 %)			
Anwendung				3.07	4	.547
EKG	48 (39 %)	318 (32 %)	366 (33 %)			
Insulin	34 (28 %)	339 (34 %)	373 (34 %)			
Koloskopie	40 (33 %)	334 (34 %)	374 (34 %)			
Alter	52 (\pm 12)	50 (\pm 17)	51 (\pm 16)	1.01	2, 363.38	.365
Geschlecht *				8.36	2	.0153
Weiblich	45 (37 %)	503 (51 %)	548 (49 %)			
Männlich	77 (63 %)	488 (49 %)	565 (51 %)			
ATAI	7.0 (\pm 2.0)	6.6 (\pm 2.0)	6.6 (\pm 2.0)	2.41	2, 2223	.09

Anmerkung: Bestimmung der Unterschiede durch unabhängige Stichproben-t-Tests oder Pearsons χ^2 -Tests.

Ergebnisse

Die mehrfaktorielle ANOVA ergab mehrere statistisch signifikante Haupteffekte und eine signifikante Interaktion (siehe Tabelle 6):

Tabelle 10: Ergebnisse der ANOVA

Variablen	df	SS	MS	F	p
Autonomie	1.00	1.08	1.08	4.01	.04*
Erklärbarkeit	1.00	0.09	0.09	0.32	0.57
Anwendung	2.00	5.09	2.54	9.46	< .001***
Gruppe	1.00	2.05	2.05	7.60	0.01*
Alter	1.00	0.60	0.60	2.22	0.14
Geschlecht	1.00	0.37	0.37	1.36	0.24
ATAI	1.00	86.71	86.71	322.35	< .001***
Autonomie X Erklärbarkeit	1.00	0.00	0.00	0.01	0.92
Autonomie X Anwendung	2.00	0.21	0.11	0.40	0.67
Erklärbarkeit X Anwendung	2.00	1.72	0.86	3.19	0.04*
Autonomie X Gruppe	1.00	0.01	0.01	0.05	0.82
Erklärbarkeit X Gruppe	1.00	0.96	0.96	3.56	0.06†
Anwendung X Gruppe	2.00	0.25	0.12	0.46	0.63
Autonomie X Erklärbarkeit X Anwendung	2.00	0.64	0.32	1.19	0.30
Autonomie X Erklärbarkeit X Gruppe	1.00	0.86	0.86	3.21	0.07†
Autonomie X Anwendung X Gruppe	2.00	0.12	0.06	0.23	0.80
Erklärbarkeit X Anwendung X Gruppe	2.00	0.36	0.18	0.67	0.51
Autonomie X Erklärbarkeit X Anwendung X Gruppe	2.00	0.69	0.34	1.27	0.28

Anmerkung: Alter, Geschlecht und ATAI wurden als Kovariaten aufgenommen; *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .1$

Haupteffekt KI-Autonomie: Es zeigte sich ein signifikanter Haupteffekt, $F(1, 1086) = 4.01$, $p = .046$, $\eta^2 p = .004$. Post-hoc-Tests (Tukey-HSD) zeigten einen signifikanten Unterschied zwischen menschlicher Entscheidung ($M = 3.22$, $SD = 0.60$) und KI-Entscheidung ($M = 3.16$, $SD = 0.59$), mit einer mittleren Differenz von -0.06 , 95% CI $[-0.12, -0.001]$, $p = .046$. Dies deutet darauf hin, dass die Akzeptanz signifikant höher ist, wenn Menschen die Entscheidungen treffen, im Vergleich zu KI-getroffenen Entscheidungen.

Haupteffekt Art der KI-Anwendung: Es wurde ein hochsignifikanter Haupteffekt festgestellt, $F(2, 1086) = 9.46$, $p < .001$, $\eta^2 p = .017$. Post-hoc-Analysen (Tukey-HSD) ergaben:

- Insulinpumpe ($M = 3.28$, $SD = 0.58$) vs. EKG ($M = 3.16$, $SD = 0.61$): signifikante Differenz von 0.13 , 95% CI $[0.04, 0.22]$, $p = .002$
- Koloskopie ($M = 3.13$, $SD = 0.59$) vs. EKG: kein signifikanter Unterschied, $p = .824$
- Koloskopie vs. Insulinpumpe: signifikante Differenz von -0.15 , 95% CI $[-0.24, -0.06]$, $p < .001$

Diese Ergebnisse zeigen, dass die Insulinpumpe die höchste Akzeptanz aufweist, gefolgt von EKG und Koloskopie.

Haupteffekt Gruppenzugehörigkeit: Es zeigte sich ein signifikanter Haupteffekt, $F(1, 1086) = 7.60$, $p = .006$, $\eta^2 p = .007$. Post-hoc-Tests zeigten, dass Versicherte ($M = 3.21$, $SD = 0.59$) einen signifikant höheren Akzeptanzwert aufwiesen als ÄrztInnen ($M = 3.06$, $SD = 0.64$), mit einer mittleren Differenz von 0.14 , 95% CI $[0.04, 0.23]$, $p = .006$. Dies deutet auf eine höhere Akzeptanz gegenüber den KI-Anwendungen seitens der Versicherten hin.

Haupteffekt ATAI: Es wurde ein hochsignifikanter Haupteffekt festgestellt, $F(1, 1086) = 322.35$, $p < .001$, $\eta^2p = .229$. Dies zeigt, dass die allgemeine Einstellung gegenüber KI (gemessen durch ATAI) einen starken Einfluss auf die Akzeptanz der spezifischen KI-Anwendungen hat.

Abbildung 9 zeigt die Haupteffekte der experimentellen Between-Subjects-Faktoren. Die Akzeptanz war signifikant höher bei menschlicher Entscheidungshoheit im Vergleich zu KI-Autonomie ($p = .046$). Bezüglich der Anwendungsarten zeigte die Insulinpumpe die höchste Akzeptanz, signifikant höher als EKG ($p = .002$) und Koloskopie ($p < .001$). Versicherte wiesen eine signifikant höhere Akzeptanz auf als ÄrztInnen ($p = .006$).

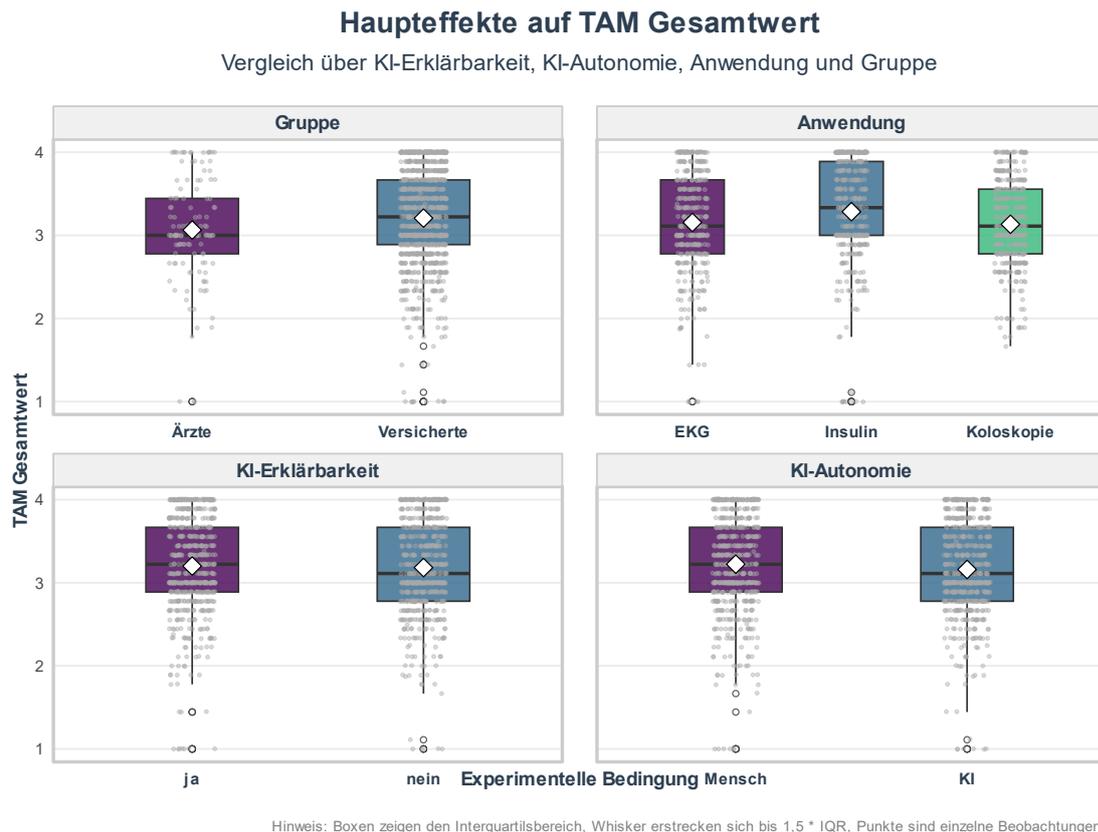


Abbildung 9: Haupteffekte der experimentellen Faktoren auf die KI-Akzeptanz.

Interaktionseffekt KI-Erklärbarkeit x Anwendung: Es wurde eine statistisch signifikante Interaktion gefunden, $F(2, 1086) = 3.19$, $p = .041$, $\eta^2p = .006$. Dies deutet darauf hin, dass der Effekt der KI-Erklärbarkeit auf den Akzeptanzwert je nach KI-Anwendung (EKG, Insulinpumpe, Koloskopie) variiert. Post-hoc-Analysen mit Tukey-Korrektur zeigten mehrere signifikante Unterschiede zwischen:

- kurzer Erklärung bei EKG und langer Erklärung bei Insulin ($p = .037$)
- kurzer Erklärung bei Insulin und langer Erklärung bei Koloskopie ($p < .001$)
- langer Erklärung bei Insulin und langer Erklärung bei Koloskopie ($p < .001$)

Diese Ergebnisse weisen auf kontextspezifische Effekte der KI-Erklärbarkeit hin. Die Insulin-Anwendung mit langer Erklärung zeigt signifikant höhere Akzeptanzwerte als die Koloskopie mit langer Erklärung. Die Insulin-Anwendung mit kurzer Erklärung wird signifikant besser akzeptiert als die Koloskopie mit langer Erklärung. Die praktische Bedeutsamkeit dieser Unterschiede sollte unter Berücksichtigung der marginalen Signifikanz der Interaktion und der kleinen Effektstärke vorsichtig interpretiert werden (siehe auch Abbildung 10).

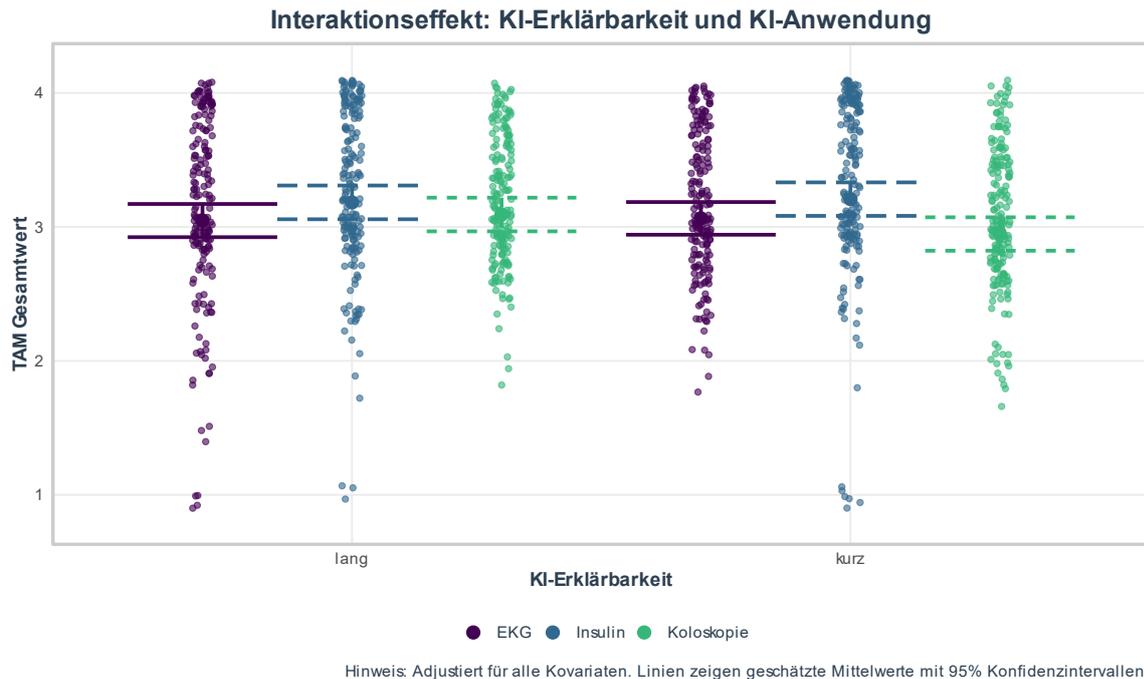


Abbildung 10: Interaktionseffekt KI-Erklärbarkeit und KI-Anwendung

Interaktion KI-Erklärbarkeit x Gruppenzugehörigkeit: Die Analyse ergab eine nicht-signifikante Interaktion zwischen KI-Erklärbarkeit und Gruppenzugehörigkeit, $F(1, 1086) = 3.56$, $p = .059$, $\eta^2p = .003$. Obwohl dieser Effekt das konventionelle Signifikanzniveau von $p < .05$ knapp verfehlt, wird er aufgrund seiner theoretischen Relevanz für das Verständnis gruppenspezifischer Informationsbedürfnisse, seiner praktischen Bedeutung für die zielgruppengerechte Gestaltung von KI-Erklärungen im Gesundheitswesen sowie der Möglichkeit eines Typ-II-Fehlers aufgrund der ungleichen Gruppengrößen zwischen ÄrztInnen und Versicherten berichtet.

Die deskriptiven Muster (siehe Abbildung 11) deuten auf möglicherweise unterschiedliche Präferenzen bezüglich der Erklärungslänge zwischen den Gruppen hin. ÄrztInnen könnten längere, detailliertere Erklärungen bevorzugen, was zu einer höheren Akzeptanz führt; Versicherte hingegen könnten eine leicht höhere Akzeptanz bei kürzeren Erklärungen zeigen.

Für eine statistisch abgesicherte Interpretation dieser Unterschiede wären jedoch weitere Studien mit größeren und ausbalancierten Stichproben erforderlich. Die praktische Bedeutsamkeit dieser tendenziellen Unterschiede sollte unter Berücksichtigung der nicht-signifikanz besonders vorsichtig interpretiert werden.

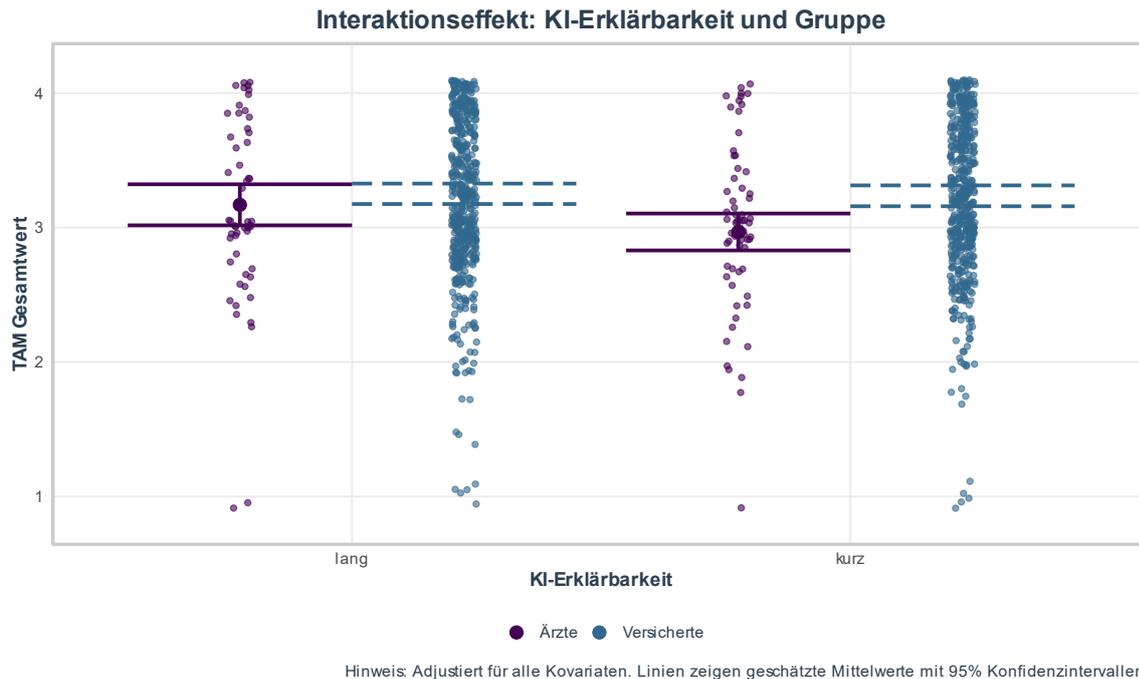


Abbildung 11: Interaktionseffekt KI-Erklärbarkeit und Gruppe

Interaktion Autonomie x Erklärbarkeit x Gruppe: Die Analyse ergab eine nicht-signifikante Dreifach-Interaktion zwischen KI-Autonomie, Erklärbarkeit und Gruppenzugehörigkeit, $F(1, 1086) = 3.21$, $p = .073$, $\eta^2 p = .003$. Obwohl dieser Effekt das konventionelle Signifikanzniveau von $p < .05$ knapp verfehlt, wurden aufgrund der theoretischen Relevanz für die zielgruppenspezifische Gestaltung von KI-Systemen simple-slope-Analysen durchgeführt (siehe Abbildung 12).

Bei ÄrztInnen zeigt sich ein differenziertes Bild in Abhängigkeit von der Autonomie-Bedingung: Unter menschlicher Entscheidungshoheit (violette Linie) ist ein signifikant negativer Slope zu beobachten ($b = -0.365$, $SE = 0.132$, $CI [-0.623, -0.107]$), der einen deutlichen Abfall der Akzeptanz von langer zu kurzer Erklärung anzeigt. Im Gegensatz dazu zeigt sich unter KI-Autonomie (gelbe Linie) ein nicht signifikanter, leicht negativer Slope ($b = -0.029$, $SE = 0.138$, $CI [-0.299, 0.241]$). Bei Versicherten ergibt sich ein deutlich homogeneres Bild mit nahezu parallelen, leicht abfallenden Linien für beide Autonomie-Bedingungen. Weder unter menschlicher Entscheidungshoheit ($b = 0.003$, $SE = 0.047$, $CI [-0.089, 0.095]$) noch unter KI-Autonomie ($b = -0.032$, $SE = 0.047$, $CI [-0.124, 0.059]$) zeigen sich signifikante Slopes. Diese Ergebnisse verdeutlichen, dass die Länge der KI-Erklärung nur bei ÄrztInnen unter der Bedingung menschlicher Autonomie einen bedeutsamen Einfluss auf die Akzeptanz hat, wobei längere Erklärungen zu deutlich höherer Akzeptanz führen. Bei Versicherten hingegen spielt die Erklärungslänge eine untergeordnete Rolle, unabhängig von der Autonomie-Bedingung.

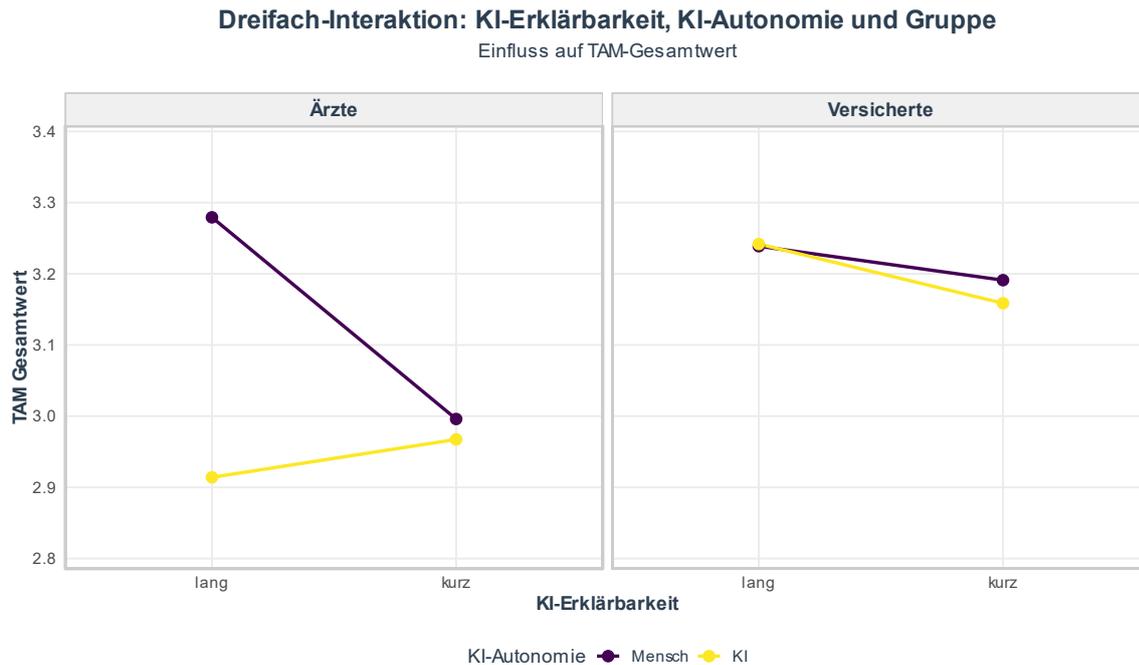


Abbildung 12: Dreifach-Interaktion KI-Erklärbarkeit, KI-Autonomie und Gruppe

6.2.2 Studie 2: KI-Akzeptanz-Skala

Im Rahmen von AP3 wurde das im Berichtszeitraum 2021 entwickelte Modell der KI-Akzeptanz operationalisiert und anhand eines online-basierten Pretests mit Blick auf die psychometrische Qualität getestet. Die Vorstudie wurde durchgeführt, um mit einem sparsamen und initial psychometrisch validierten Instrument in die Hauptstudie gehen zu können. Die Ziele dieser Vorstudie waren:

- Die empirische Überprüfung der theoretisch postulierten Faktorenstruktur
- Die Identifikation der psychometrisch stärksten Items pro Dimension
- Die Entwicklung einer ökonomischen Kurzversion des Instruments
- Die Sicherstellung der Messinvarianz zwischen verschiedenen Zielgruppen
- Die initiale Validierung des übergeordneten Akzeptanzmodells

Hierzu wurden 191 Versicherte sowie 139 ÄrztInnen und Medizinstudierende gebeten, eine Vielzahl von Items zur KI-Akzeptanz zu beantworten. Das resultierende Instrument wurde im Rahmen von AP5 mit Probanden aus der Hauptstudie validiert. Durch die unterschiedlichen Rekrutierungsstrategien kann eine Doppelteilnahme von Probanden mit großer Wahrscheinlichkeit ausgeschlossen werden. Die Ziele der Hauptstudie waren:

- Die Kreuzvalidierung der Faktorenstruktur an einer unabhängigen Stichprobe
- Die Überprüfung der Konstruktvalidität des entwickelten Messinstruments
- Die Untersuchung von Gruppenunterschieden in der KI-Akzeptanz
- Die Analyse von Einflussfaktoren auf die Nutzungsbereitschaft von KI im Gesundheitswesen
- Vergleich der neuen Skala mit bestehenden Maßen zur KI-Akzeptanz (ATAI)

Um für mögliche Verzerrungen zu kontrollieren, wurden zusätzliche Variablen wie Geschlecht, Bildung und subjektive Gesundheit in die Analysen einbezogen. Durch diesen Ansatz sollten statistisch signifikante Unterschiede in der Akzeptanz von KI-Anwendungen zwischen verschiedenen Gruppen identifiziert und Faktoren ermittelt werden, die die Akzeptanz und

potenzielle Implementierung von KI-basierten Innovationen im Gesundheitswesen beeinflussen.

Im Folgenden werden die zentralen Informationen und Ergebnisse beider Studien getrennt dargestellt.

Stichprobe und Variablen - Vorstudie

Stichprobe 1 (Versicherte): Die initiale Stichprobe umfasste 207 Teilnehmende. Im Rahmen der Datenbereinigung wurden 16 Teilnehmende ausgeschlossen. Ausschlussgründe waren vorzeitiger Abbruch des Fragebogens, Nichterfüllung der Teilnahmevoraussetzungen (z.B. Minderjährigkeit) sowie Antworten in Response-Sets (gleichförmiges Ankreuzverhalten trotz invertierter Items). Die finale Stichprobe bestand aus 191 Leistungsempfängern im Alter von 20 bis 88 Jahren ($M = 54,4$; $SD = 19,5$). Die Rekrutierung erfolgte über die Verbreitung des Umfragelinks durch lokale Community-Netzwerke in Deutschland. Der Frauenanteil lag bei 54,5%, 84,3% der Teilnehmenden verfügten über Abitur oder einen höheren Bildungsabschluss. Insgesamt schätzten sich 70,3% als eher kompetent oder sehr kompetent im Umgang mit Technologien ein. Die Studie wurde als Online-Befragung durchgeführt.

Stichprobe 2 (Leistungserbringer): Die initiale Stichprobe umfasste 144 Teilnehmende. Nach Anwendung identischer Ausschlusskriterien verblieben 138 Gesundheitsdienstleister in der finalen Stichprobe. Davon waren 59,4% Medizinstudierende oder ÄrztInnen in Weiterbildung, die übrigen waren approbierte ÄrztInnen. Die Rekrutierung erfolgte über Fachschaften verschiedener deutscher Universitäten sowie über Serien-E-Mails an praktizierende ÄrztInnen. Die Altersspanne der Teilnehmenden lag zwischen 19 und 85 Jahren ($M = 32,2$; $SD = 15,5$), der Frauenanteil betrug 65,6%. Die Medizinstudierenden verteilten sich nahezu gleichmäßig auf die erste und zweite Studienhälfte. Die ÄrztInnen waren überwiegend in der Inneren Medizin und als Allgemeinmediziner tätig. 58,3% der ÄrztInnen arbeiteten in eigener Praxis, gefolgt von 27,8% in Wissenschaft oder außerhalb der Praxis. Unter den Leistungserbringern schätzten 72,9% ihre Technologiekompetenz als eher hoch oder sehr hoch ein. Die Studie wurde als Online-Befragung durchgeführt.

Ergebnisse – Vorstudie

Skalenentwicklung: Basierend auf einer Literaturanalyse wurden vier theoretische Dimensionen der KI-Akzeptanz identifiziert und operationalisiert. Jede Dimension wurde durch 11-12 Items erfasst, die mit der Einleitung "KI in der Gesundheitsversorgung..." oder "Durch den Einsatz von KI in der Gesundheitsversorgung..." begannen. Die Beantwortung erfolgte auf einer 5-stufigen Likert-Skala von "trifft überhaupt nicht zu" bis "trifft sehr zu". Die aus der Literaturanalyse abgeleiteten Dimensionen umfassten 1) KI-Nützlichkeit (12 Items) zur Messung der wahrgenommenen Vorteile der KI-Implementierung und deren direkte Auswirkungen auf das Gesundheitswesen; 2) KI-Kontrolle (11 Items) zur Erfassung des Ausmaßes der wahrgenommenen Interventionsmöglichkeiten in medizinische Prozesse unter Berücksichtigung von Verantwortlichkeit und Rechenschaftspflicht; 3) KI-Angst (11 Items) zur Erfassung negativer emotionaler Reaktionen gegenüber KI; 4) KI-Vertrauen zur Messung kognitiver Vertrauensaspekte, z.B. bezogen auf Aspekte der Datensicherheit und Zuverlässigkeit medizinischer Diagnostik. Die Items wurden interdisziplinär von einem Forschungsteam mit Expertise in Psychologie, Gerontologie und Medizin entwickelt. Um unverfälschte Meinungen zu erfassen, wurde bewusst auf eine vorherige Definition von KI verzichtet. Nach jeder Dimension wurden drei kurze Fragen zur Bewertung der Itemqualität gestellt ("Ich finde die obigen Fragen kompliziert/verständlich/schwer zu beantworten").

Faktorielle Validität: hypothetisierte Vier-Faktoren-Struktur wurde durch vergleichende Modellanalysen bestätigt. Modelle mit zwei und drei Faktoren wiesen höhere RMSEA- und niedrigere TLI-Werte auf als das Vier-Faktoren-Modell. Ein Fünf-Faktoren-Modell konvergierte nicht. Das Vier-Faktoren-Modell zeigte akzeptable Fit-Indizes (RMSEA = 0,057; TLI = 0,890).

Zur weiteren Validierung wurden explorative Faktorenanalysen (EFA) mit Maximum-Likelihood-Schätzung und Oblimin-Rotation durchgeführt. Um die Vergleichbarkeit zwischen Patienten und Leistungserbringern zu gewährleisten, erfolgte die EFA simultan für beide Stichproben. Items wurden nur dann ausgeschlossen, wenn in beiden Gruppen mindestens eines der folgenden Kriterien erfüllt war:

- Niedrigste Ladung auf dem zugehörigen Faktor
- Generell geringe Faktorladung ($< .30$)
- Ähnliche Kreuzladungen auf mehreren Faktoren

Die ausgeschlossenen Items wären in beiden Stichproben unabhängig vom gewählten Vorgehen eliminiert worden, was auf eine hohe Übereinstimmung zwischen Patienten und Leistungserbringern hindeutet. Die finale Skala bestätigt die theoretisch postulierten vier Dimensionen:

- KI-Nützlichkeit
- KI-Kontrolle
- KI-Angst
- KI-Vertrauen

Jede Dimension wird durch fünf Items repräsentiert. Das finale Modell zeigt sehr gute Anpassungswerte (RMSEA = 0,039; TLI = 0,976) mit einer kumulierten erklärten Varianz von 56%. Die theoretisch abgeleiteten Dimensionen wurden durch die Faktorenanalysen empirisch bestätigt, auch ohne Vorgabe einer Vier-Faktoren-Lösung.

Messinvarianz: Zur Absicherung der Vergleichbarkeit zwischen den Stichproben wurde eine Messinvarianzprüfung nach einem stufenweisen Ansatz durchgeführt. Der Fokus lag dabei auf der Invarianz der Faktorstruktur und Itemladungen mittels Vergleichs zweier Modelle:

- Modell 1: Ohne Invarianzbeschränkungen
- Modell 2: Mit Gleichheitsbeschränkungen der Faktorladungen

Die Analysen bestätigten metrische Invarianz für die Gruppenvariable:

- $\chi^2 (16, 392) = 488,35$; $p = 0,232$
- CFI = 0,960
- RMSEA = 0,055

Sowie für die Altersgruppen (< 60 Jahre vs. ≥ 60 Jahre):

- $\chi^2 (16, 392) = 496,56$; $p = 0,920$
- CFI = 0,957
- RMSEA = 0,057

Die CFI-Differenzen lagen durchgängig unter 0,01, was die Anwendbarkeit der Skala sowohl für Leistungsempfänger als auch Leistungserbringer über die untersuchten Altersgruppen hinweg bestätigt.

Prädiktive Validität: Die Strukturgleichungsmodellierung der KI-Akzeptanz ergab ein Modell höherer Ordnung, das substantielle Zusammenhänge aller Faktoren erster Ordnung mit einem gemeinsamen übergeordneten Faktor der KI-Akzeptanz aufwies. Zur umfassenden Analyse wurden verschiedene Kovariaten erhoben. Die wahrgenommene Technologiekompetenz (PTC) wurde mittels dreier Items auf einer 5-stufigen Skala erfasst: "Generell nutze ich moderne Technologie häufig", "Ich interessiere mich für technologische Interventionen" und "Ich halte mich für kompetent genug im Umgang mit moderner Technologie". Diese Skala zeigte eine zufriedenstellende interne Konsistenz ($\alpha = 0,78$). Die Verhaltensintention (BI) wurde durch das Einzelitem "Ich würde KI-basierte Technologie nutzen" auf einer 5-stufigen Skala gemessen. Zusätzlich wurde der selbsteingeschätzte Gesundheitszustand auf einer 5-stufigen Skala von "schlecht" bis "ausgezeichnet" erhoben, sowie soziodemografische

Variablen wie Alter in Jahren und Geschlecht. Die Kombination des Modells höherer Ordnung mit einem strukturellen Regressionsmodell zur Vorhersage der Verhaltensintention unter Kontrolle der Kovariaten zeigte differenzierte Effekte. Der übergeordnete Faktor der KI-Akzeptanz erwies sich als starker Prädiktor der Verhaltensintention ($b = 0,85$; $\beta = 0,67$; $p < 0,001$), ebenso wie die wahrgenommene Technologiekompetenz ($b = 0,29$; $\beta = 0,23$; $p < 0,001$). Demgegenüber zeigten sich keine signifikanten Effekte für Geschlecht ($b = 0,08$; $\beta = 0,04$; $p = 0,400$), Gesundheitszustand ($b = -0,07$; $\beta = -0,07$; $p = 0,161$), Alter ($b = -0,00$; $\beta = -0,00$; $p = 0,977$) oder Gruppenzugehörigkeit ($b = 0,05$; $\beta = 0,03$; $p = 0,578$). Diese Ergebnisse verdeutlichen, dass eine höhere KI-Akzeptanz mit einer stärkeren Verhaltensintention einhergeht, wobei keine Unterschiede zwischen Patienten und Leistungserbringern oder verschiedenen Altersgruppen bestehen. Die wahrgenommene Technologiekompetenz stellt sich als bedeutsamer zusätzlicher Einflussfaktor auf die Verhaltensintention heraus.

Stichprobe und Variablen - Hauptstudie

Die Stichprobe dieser Studie umfasste insgesamt 1,259 Teilnehmer, davon 157 ÄrztInnen und 1,102 Versicherte. Das durchschnittliche Alter der Gesamtstichprobe betrug 51 Jahre ($SD = 17$), wobei sich kein signifikanter Altersunterschied zwischen ÄrztInnen und Versicherten zeigte, $t(2, 463.73) = 0.72$, $p = .489$. Die Geschlechterverteilung unterschied sich signifikant zwischen den Gruppen, $\chi^2(2) = 12.83$, $p = .0016$, mit einem höheren Anteil männlicher Teilnehmer unter den ÄrztInnen (64%) im Vergleich zu den Versicherten (48%). Die subjektive Gesundheit wurde mit einer Einzelfrage erfasst ("Wie würden Sie Ihren gegenwärtigen Gesundheitszustand beschreiben?") und auf einer 5-Punkt-Skala von 1 (schlecht) bis 5 (sehr gut) bewertet. ÄrztInnen berichteten eine signifikant bessere subjektive Gesundheit ($M = 4.0$, $SD = 0.75$) als Versicherte ($M = 3.7$, $SD = 0.93$), $t(2, 2185) = 7.39$, $p < .001$. Die generelle Einstellung zu Künstlicher Intelligenz (KI) wurde mit dem ATAI (Attitude Towards Artificial Intelligence) erfasst, einer Skala von 0 (starke Ablehnung) bis 10 (starke Zustimmung), wobei sich signifikante Unterschiede zwischen ÄrztInnen und Versicherten zeigten, $t(2, 2339) = 3.82$, $p = .022$. Zusätzlich wurden im Rahmen dieser Studie eigene Skalen zur spezifischen KI-Akzeptanz entwickelt: KI-Vertrauen, KI-Angst, KI-Kontrolle und KI-Nützlichkeit, die jeweils als Mittelwerte der entsprechenden Items gebildet wurden. Diese Skalen wurden auf einer 5-Punkt-Likert-Skala von 1 (trifft überhaupt nicht zu) bis 5 (trifft genau zu) gemessen. Der KI-Gesamtwert wurde als Durchschnitt dieser vier Skalen berechnet. Signifikante Unterschiede zwischen ÄrztInnen und Versicherten zeigten sich insbesondere bei der wahrgenommenen Kontrolle über KI, $t(2, 2339) = 11.26$, $p < .001$. Die Technikakzeptanz (TAM-Gesamtwert), die die Beurteilung der präsentierten KI-Stimuli darstellt, wurde auf einer 4-Punkt-Skala von 1 (Trifft gar nicht zu) bis 4 (Trifft genau zu) gemessen und wies ebenfalls signifikante Unterschiede zwischen den Gruppen auf, $t(2, 2225) = 3.15$, $p = .0428$. Fehlende Werte auf den Variablen wurden mittels multipler Imputation unter Verwendung der Predictive Mean Matching (PMM) Methode mit fünf Imputationen und 50 Iterationen geschätzt, um eine vollständige Datenmatrix für die Analysen zu erhalten. Tabelle 7 zeigt eine detaillierte Deskription.

Tabelle 11: Deskription

Variablen	ÄrztInnen (N = 157)	Versicherte (N = 1102)	Gesamt (N = 1259)	t / χ^2	df	p
Alter	52 (± 13)	50 (± 17)	51 (± 17)	0.72	2, 463.73	.489
Geschlecht **				12.83	2	.0016
Weiblich	57 (36 %)	568 (52 %)	625 (50 %)			
Männlich	100 (64 %)	533 (48 %)	633 (50 %)			
ATAI *	7.1 (± 2.0)	6.6 (± 2.0)	6.6 (± 2.0)	3.82	2, 2339	.022
KI-Vertrauen	3.4 (± 0.68)	3.5 (± 0.81)	3.5 (± 0.79)	1.14	2, 2341	.319
KI-Angst	4.0 (± 0.98)	3.9 (± 1.0)	3.9 (± 1.0)	0.53	2, 2337	.591
KI-Kontrolle ***	3.5 (± 0.84)	3.1 (± 0.84)	3.2 (± 0.84)	11.26	2, 2339	< .001
KI-Nützlichkeit	3.6 (± 0.83)	3.8 (± 0.79)	3.8 (± 0.79)	2.42	2, 2337	.0888
KI-Gesamtwert	3.6 (± 0.70)	3.6 (± 0.70)	3.6 (± 0.70)	0.28	2, 2341	.758
Subj. Gesundheit ***	4.0 (± 0.75)	3.7 (± 0.93)	3.7 (± 0.92)	7.39	2, 2185	< .001
TAM-Gesamtwert *	3.1 (± 0.64)	3.2 (± 0.59)	3.2 (± 0.60)	3.15	2, 2225	.0428

Anmerkung: Bestimmung der Unterschiede durch unabhängige Stichproben-t-Tests oder Pearsons χ^2 -Tests.

Ergebnisse

Faktorielle Validität: Zur Untersuchung der zugrundeliegenden Struktur der Einstellungen gegenüber künstlicher Intelligenz (KI) in der Gesundheitsversorgung wurde eine explorative Faktorenanalyse (EFA) durchgeführt. Sowohl der Scree-Test als auch der Paralleltest deuteten auf eine Vier-Faktor-Lösung hin. Basierend auf diesen Ergebnissen wurde eine EFA mit vier Faktoren unter Verwendung der Minimum-Residuen-Methode (MINRES) und Oblimin-Rotation durchgeführt.

Die Analyse ergab eine Vier-Faktor-Lösung, die 68% der Gesamtvarianz erklärte. Die extrahierten Faktoren wurden entsprechend der theoretischen Einbettung als „Angst“, „Nutzen“, „Kontrolle“ und „Vertrauen“ interpretiert. Die Ladungsstruktur ist in Tabelle 8 dargestellt.

Der erste Faktor, „Angst“, umfasste fünf Items mit Faktorladungen zwischen 0.74 und 0.95. Diese Items beschreiben negative emotionale Reaktionen auf KI in der Gesundheitsversorgung, wie z.B. „fühlt sich für mich bedrohlich an“ (Faktorladung = 0.95) und „ist beklemmend für mich“ (Faktorladung = 0.95). Die Kommunalitäten (h^2) für diese Items lagen zwischen 0.70 und 0.87, was auf eine gute Erklärung der Varianz durch den extrahierten Faktor hindeutet.

Der zweite Faktor, „Nutzen“, bestand aus fünf Items mit Faktorladungen zwischen 0.73 und 0.92. Diese Items beziehen sich auf die wahrgenommenen Vorteile von KI in der Gesundheitsversorgung, wie „können Krankheiten besser behandelt werden“ (Faktorladung = 0.92) und „wird die Gesundheitsversorgung von Patienten verbessert“ (Faktorladung = 0.87). Die Kommunalitäten für diese Items lagen zwischen 0.62 und 0.81.

Der dritte Faktor, „Kontrolle“, umfasste fünf Items mit Faktorladungen zwischen 0.59 und 0.83. Diese Items beziehen sich auf die wahrgenommene Kontrolle über medizinische Entscheidungen bei der Verwendung von KI, wie z.B. „unterliegen medizinische Entscheidungen meiner Kontrolle“ (Faktorladung = 0.83). Die Kommunalitäten für diese Items lagen zwischen 0.49 und 0.66.

Der vierte Faktor, „Vertrauen“, bestand aus fünf Items mit Faktorladungen zwischen 0.51 und 0.84. Diese Items reflektieren das Vertrauen in KI-basierte medizinische Empfehlungen und Diagnosen, wie "verringert mein Vertrauen in medizinische Empfehlungen" (Faktorladung = 0.84, negativ formuliert). Die Kommunalitäten für diese Items lagen zwischen 0.50 und 0.72.

Insgesamt zeigten die Items gute psychometrische Eigenschaften mit Faktorladungen über 0.50 und Kommunalitäten größtenteils über 0.60. Die Vier-Faktor-Struktur, die durch den Scree-Test und Paralleltest unterstützt und durch die MINRES-Methode mit Oblimin-Rotation extrahiert wurde, bietet eine differenzierte Sicht auf die Einstellungen gegenüber KI in der Gesundheitsversorgung, die emotionale Reaktionen (Angst), wahrgenommene Vorteile (Nutzen), Kontrollüberzeugungen und Vertrauensaspekte umfasst.

Die psychometrische Analyse der vier Dimensionen des KI-Akzeptanz-Instruments zeigte durchweg sehr gute interne Konsistenzen. Die Subskala "KI-Angst" wies mit $\alpha = 0,95$ die höchste Reliabilität auf, gefolgt von der Dimension "KI-Nutzen" mit $\alpha = 0,93$. Die Skalen "KI-Vertrauen" und "KI-Kontrolle" zeigten mit jeweils $\alpha = 0,87$ ebenfalls sehr zufriedenstellende Werte. Diese durchgängig hohen Reliabilitätskoeffizienten bestätigen die psychometrische Qualität des entwickelten Instruments und seine Eignung für den Einsatz in der Forschungspraxis.

Die Analyse der Faktorkorrelationen zeigte moderate bis starke Zusammenhänge zwischen den extrahierten Faktoren:

- Angst und Nutzen: $r = 0.57$
- Angst und Kontrolle: $r = 0.36$
- Angst und Vertrauen: $r = 0.71$
- Nutzen und Kontrolle: $r = 0.53$
- Nutzen und Vertrauen: $r = 0.72$
- Kontrolle und Vertrauen: $r = 0.45$

Diese Korrelationen deuten darauf hin, dass die vier Faktoren, obwohl sie distinkte Aspekte der Einstellungen gegenüber KI in der Gesundheitsversorgung erfassen, nicht vollständig unabhängig voneinander sind. Insbesondere die starken Korrelationen zwischen Vertrauen und den anderen Faktoren (Angst: $r = 0.71$, Nutzen: $r = 0.72$) sowie zwischen Angst und Nutzen ($r = 0.57$) legen nahe, dass diese Faktoren möglicherweise Teil eines übergeordneten Konstrukts sind. Dies impliziert, dass neben der differenzierten Betrachtung der einzelnen Dimensionen auch die Bildung eines Gesamtwertes gerechtfertigt ist. Dieser Gesamtwert stellt eine allgemeine Einstellung oder Akzeptanz gegenüber KI in der Gesundheitsversorgung dar, die sich aus den entsprechenden Subskalen zusammensetzt.

Tabelle 12: Darstellung der Ladungsstruktur

Aussage	Item	Angst	Nutzen	Kontrolle	Vertrauen	h ²
KI in der Gesundheitsversorgung	erhöht mein Vertrauen in medizinische Empfehlungen.				0.75	0.72
	verringert mein Vertrauen in medizinische Empfehlungen.				0.84	0.69
	finde ich zuverlässig.				0.53	0.52
	verringert mein Vertrauen in medizinische Diagnosen.				0.63	0.50
	finde ich unsicher.				0.51	0.59
KI in der Gesundheitsversorgung	erfüllt mich mit Angst.	0.85				0.78
	bereitet mir Unbehagen.	0.83				0.83
	fühlt sich für mich bedrohlich an.	0.95				0.85
	ist beklemmend für mich.	0.95				0.87
	lässt mich sorgenvoll in die Zukunft blicken.	0.74				0.70
Durch den Einsatz von KI in der Gesundheitsversorgung	unterliegen medizinische Entscheidungen meiner Kontrolle.			0.83		0.63
	kann ich medizinische Entscheidungen hinterfragen.			0.66		0.50
	trifft der Mensch die wichtigen medizinischen Entscheidungen.			0.59		0.49
	unterliegt der Umgang mit medizinischen Entscheidungen meiner Kontrolle.			0.8		0.66
	habe ich Einfluss auf medizinische Entscheidungen.			0.77		0.61
Durch den Einsatz von KI in der Gesundheitsversorgung	profitieren Patienten von maßgeschneiderten Therapien.		0.73			0.62
	wird die Gesundheitsversorgung von Patienten verbessert.		0.87			0.81
	können Krankheiten besser behandelt werden.		0.92			0.78
	profitiert das Gesundheitssystem insgesamt.		0.77			0.64
	werden medizinische Empfehlungen für Menschen verbessert.		0.81			0.75

Messinvarianz: Zur Überprüfung der Messinvarianz des Vier-Faktoren-Modells der KI-Akzeptanz zwischen ÄrztInnen und Versicherten wurde eine schrittweise Analyse der konfiguralen, metrischen und skalaren Invarianz durchgeführt.

Das konfigurale Modell, welches die gleiche Faktorstruktur in beiden Gruppen annimmt, zeigte einen akzeptablen Fit ($\chi^2(328) = 1160.010$, $p < .001$, CFI = .956, TLI = .949, RMSEA = .066 [.062, .070], SRMR = .043). Dies deutet darauf hin, dass die grundlegende Faktorstruktur in beiden Gruppen ähnlich ist.

Im nächsten Schritt wurde die metrische Invarianz getestet, indem die Faktorladungen über die Gruppen hinweg gleichgesetzt wurden. Der Modellvergleich zeigte eine signifikante Verschlechterung des Modell-Fits ($\Delta\chi^2(16) = 32.607$, $p = .008$), wobei die Veränderungen in den Fit-Indizes minimal waren ($\Delta CFI = .001$, $\Delta RMSEA = .001$). Trotz der statistischen Signifikanz kann aufgrund der geringen praktischen Bedeutsamkeit der Unterschiede von einer partiellen metrischen Invarianz ausgegangen werden.

Schließlich wurde die skalare Invarianz getestet, indem zusätzlich zu den Faktorladungen auch die Intercepts über die Gruppen hinweg gleichgesetzt wurden. Auch hier zeigte der Modellvergleich eine signifikante Verschlechterung des Fits ($\Delta\chi^2(16) = 29.831$, $p = .019$), wobei die Veränderungen in den Fit-Indizes erneut minimal waren ($\Delta CFI = .001$, $\Delta RMSEA = .001$).

Insgesamt deuten die Ergebnisse auf eine akzeptable konfigurale Invarianz und eine partielle metrische und skalare Invarianz hin. Obwohl die χ^2 -Differenztests signifikant waren, waren die Veränderungen in den praktischen Fit-Indizes vernachlässigbar klein. Dies legt nahe, dass das Messmodell der KI-Akzeptanz für ÄrztInnen und Versicherte weitgehend vergleichbar ist. Für die praktische Anwendung kann das Modell als ausreichend invariant betrachtet werden, um Vergleiche zwischen den Gruppen zu ermöglichen.

Prädiktive Validität: Zur Untersuchung der prädiktiven Validität des entwickelten Messmodells für KI-Akzeptanz wurde ein Strukturgleichungsmodell berechnet (siehe Abbildung 13). Dieses Modell integrierte das Vier-Faktoren-Messmodell der KI-Akzeptanz (Vertrauen, Emotion, Kontrolle und Nutzen) als Prädiktoren für die allgemeine Technikakzeptanz (TAM-Gesamtwert), unter Kontrolle relevanter soziodemografischer und gesundheitsbezogener Variablen.

Das Gesamtmodell zeigte einen akzeptablen Fit ($\chi^2(285) = 2267.036$, $p < .001$, $CFI = .897$, $TLI = .882$, $RMSEA = .080$ [.077, .083], $SRMR = .142$). Obwohl einige Fit-Indizes unter den empfohlenen Schwellenwerten liegen, kann das Modell aufgrund der Komplexität und Stichprobengröße als angemessen betrachtet werden.

Die Ergebnisse zeigen, dass der latente Faktor zweiter Ordnung „KI-Akzeptanz“, der die vier Dimensionen der KI-Akzeptanz zusammenfasst, ein starker und signifikanter Prädiktor für den TAM-Gesamtwert ist ($\beta = .559$, $p < .001$). Dies deutet auf eine hohe prädiktive Validität des entwickelten Messmodells für die allgemeine Akzeptanz gegenüber den experimentellen Stimuli hin.

Unter den Kontrollvariablen erwiesen sich das Alter ($\beta = -.057$, $p = .032$) und die allgemeine Einstellung zu KI (ATAI) ($\beta = .102$, $p < .001$) als signifikante Prädiktoren für den TAM-Gesamtwert. Ein marginal signifikanter Effekt wurde für die Gruppenzugehörigkeit beobachtet (ÄrztInnen vs. Versicherte) ($\beta = .050$, $p = .051$), was auf potenzielle Unterschiede in der Technikakzeptanz zwischen diesen Gruppen hindeutet.

Geschlecht ($\beta = -.050$, $p = .060$) und subjektive Gesundheit ($\beta = -.005$, $p = .851$) zeigten keine signifikanten Effekte auf den TAM-Gesamtwert.

Insgesamt erklärt das Modell 33% der Varianz im TAM-Gesamtwert ($R^2 = 1 - .670 = .330$), was auf eine substantielle Erklärungskraft des entwickelten Messmodells für KI-Akzeptanz hinweist. Diese Ergebnisse unterstützen die prädiktive Validität des Vier-Faktoren-Modells der KI-Akzeptanz und zeigen seine Relevanz für die Vorhersage der Akzeptanz gegenüber konkreten KI-Lösungen im Gesundheitskontext, auch unter Berücksichtigung relevanter demografischer und gesundheitsbezogener Faktoren.

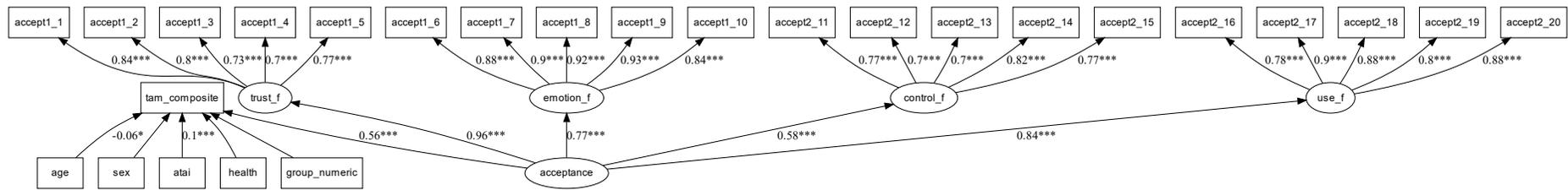


Abbildung 13: Strukturgleichungsmodell zur Messung von KI-Akzeptanz

6.2.3 Studie 3: Prädiktoren der KI-Akzeptanz

Ein weiterer Befund betrifft die Analyse potenzieller Einflussfaktoren auf die KI-Akzeptanz, gemessen am TAM-Gesamtwert. Hierfür wurden drei verschiedenen regularisierte Regressionsverfahren aus dem Bereich des maschinellen Lernens eingesetzt: LASSO (Least Absolute Shrinkage and Selection Operator), Elastic Net und Square-root LASSO. Diese Methoden wurden gewählt, um das Problem der Multikollinearität zu adressieren und ein sparsames Modell zu erhalten, das die wichtigsten Prädiktoren identifiziert. Zusätzlich wurde ein Random Forest als nicht-parametrisches Ensemble-Lernverfahren eingesetzt, um nicht-lineare Beziehungen und Interaktionen zwischen den Variablen zu erfassen. Zur Validierung der Modelle wurde eine 5-fache Kreuzvalidierung durchgeführt. Diese Verfahren eignen sich gut für hochdimensionale Datensätze und vermeiden Überanpassung.

Stichprobe und Variablen – Versicherte

Die Studie umfasste initial 1.102 Teilnehmer aus der Gruppe der Versicherten. Nach Bereinigung fehlender Daten durch vollständige Fallanalyse bestand die endgültige analytische Stichprobe aus 950 Teilnehmern, was 86,2% der ursprünglichen Stichprobe entspricht. Die Analyse konzentrierte sich auf die Vorhersage des TAM-Gesamtwerts unter Berücksichtigung aller verfügbaren 42 Variablen in der Versichertenstudie. Dieses breite Spektrum an Variablen kann in sieben Hauptkategorien unterteilt werden: **1) Individuelle KI-Akzeptanzskalen** umfassen die in Studie 2 (siehe 6.2.2.) entwickelten Instrumente zu Vertrauen, emotionalen Reaktionen, wahrgenommener Kontrolle und Nützlichkeit von KI sowie das ATAI-Instrument zur Erfassung allgemeiner Einstellungen gegenüber KI; **2) Technologiemoivation** umfasst verschiedene individuelle motivationale Komponenten im allgemeinen Umgang mit Technologie; **3) Persönlichkeitsmerkmale** umfassen die Big Five sowie zusätzliche Maße wie Zeitperspektive, Neugier oder erlebte Obsoleszenz; **4) Gesundheitsstatus und Wohlbefinden** wurden durch subjektive und objektive Gesundheitsmaße, Gesundheitsversorgungsnutzung und Versicherungsstatus abgebildet; **5) Soziale Beziehungen und Unterstützung** beinhalteten Indikatoren der Einsamkeit, Lebenszufriedenheit und auch die Breite der Internetnutzung; **6) Experimentelle Bedingungen** sind die aus Studie 1 (siehe 6.2.1.) bekannten experimentellen Faktoren wie KI-Autonomie, KI-Erklärbarkeit, sowie Art der medizinischen Anwendung; **7) Soziodemographische Daten** umfassen neben Grundinformationen wie Alter, Geschlecht und Bildung auch detaillierte Angaben zu Beschäftigung, Wohnsituation und Infrastrukturzugang. Diese umfangreiche Variablenauswahl ermöglicht eine vielschichtige Analyse der Faktoren, die die KI-Akzeptanz im Gesundheitswesen beeinflussen.

Ergebnisse – Versicherte

Die Kreuzvalidierung ergab einen durchschnittlichen R^2 -Wert von 0,361 (95% KI: 0,351-0,371), was auf eine moderate Erklärungskraft der Modelle hindeutet. Der RMSE von 0,466 und MAE von 0,374 zeigen, dass die Modelle im Durchschnitt relativ präzise Vorhersagen treffen. Bemerkenswert ist die Konsistenz der Ergebnisse über die verschiedenen Methoden hinweg. Die regularisierten Regressionsmodelle identifizierten übereinstimmend folgende Hauptprädiktoren: KI-Nutzen ($\beta = 0,161 - 0,285$, $p < 0,001$), KI-Vertrauen ($\beta = 0,099 - 0,175$, $p < 0,01$), Technikmotivation Gesamt ($\beta = 0,101 - 0,166$, $p < 0,01$) und KI-Anwendung: Insulin ($\beta = 0,110 - 0,115$, $p < 0,01$ in LASSO und Elastic Net). Die Konsistenz dieser Ergebnisse über verschiedene Modellierungsansätze hinweg ist ein Indikator für die Robustheit der Befunde.

Die Random-Forest-Analyse bestätigte die Bedeutung der in den Regressionsmodellen identifizierten Hauptfaktoren, zeigte aber auch die Relevanz zusätzlicher Variablen auf. Die Variablenwichtigkeiten (Feature Importance) wurden relativ zum wichtigsten Prädiktor (Score 100) skaliert. Konfidenzintervalle wurden mittels Bootstrap-Resampling ermittelt. Der wahrgenommene Nutzen von KI (KI-Nutzen) erwies sich als wichtigster Prädiktor mit einem Score von 100. Darauf folgten das Vertrauen in KI (KI-Vertrauen) und die allgemeine

Technikmotivation mit Scores von etwa 43. Emotionale Reaktionen auf KI (KI-Emotion) zeigten ebenfalls eine hohe Bedeutung (Score: 39,176). Auch Faktoren wie die EKG-Anwendung und das subjektive Alterserleben spielten eine wichtige Rolle, gefolgt von allgemeinen Einstellungen zu KI (ATAI) und wahrgenommener Kontrolle über KI-Systeme. Demografische Variablen wie Alter rangierten im mittleren Bereich der Wichtigkeit, während soziale Faktoren wie soziale Beziehungen und Lebenszufriedenheit relativ gesehen weiter unten angesiedelt waren.

In der Gesamtschau erwiesen sich der wahrgenommene Nutzen (KI-Nutzen), das Vertrauen in KI (KI-Vertrauen) und die allgemeine Technologiemitivation konsistent als stärkste Prädiktoren über verschiedene analytische Ansätze hinweg. Die Random Forest Analyse berücksichtigte potenzielle nicht-lineare Beziehungen und Interaktionen, insbesondere bei demografischen Variablen, während LASSO und Elastic Net eine ähnliche Variablenselektion boten, was auf eine robuste Merkmalswichtigkeit hindeutet. Die Square-root LASSO-Methode identifizierte konservativer weniger signifikante Prädiktoren. Bemerkenswert ist die durchweg geringe Bedeutung demografischer Variablen in allen Methoden. Diese methodenübergreifende Konvergenz impliziert, dass Strategien zur Förderung der KI-Akzeptanz im Gesundheitswesen bei Versicherten primär auf die Verdeutlichung des Nutzens, die Stärkung des Vertrauens in KI und die Förderung der allgemeinen Technologiemitivation abzielen sollten. Die Diskrepanzen zwischen linearen Modellen und Random Forest, insbesondere bei Variablen wie KI-Emotion und subjektivem Alterserleben, deuten auf mögliche nicht-lineare Beziehungen oder Interaktionseffekte hin, die in zukünftigen Studien näher untersucht werden sollten.

Stichprobe und Variablen – ÄrztInnen

Die ÄrztInnen-Stichprobe dieser Studie umfasste initial 157 Teilnehmer. Nach Bereinigung fehlender Daten durch vollständige Fallanalyse bestand die endgültige analytische Stichprobe aus 116 Teilnehmern (73,9% der ursprünglichen Stichprobe). Diese begrenzte Größe muss bei der Interpretation der Ergebnisse berücksichtigt werden. Die Analyse umfasste 39 Variablen, die in sieben Kategorien eingeteilt werden können: **1) KI-Akzeptanzskalen** umfassen die in Studie 2 (siehe 6.2.2.) entwickelten Instrumente zu Vertrauen, emotionalen Reaktionen, wahrgenommener Kontrolle und Nützlichkeit von KI sowie das ATAI-Instrument zur Erfassung allgemeiner Einstellungen gegenüber KI; **2) Technologiemitivation** umfasst verschiedene individuelle motivationale Komponenten im allgemeinen Umgang mit Technologie; **3) Berufsspezifische Faktoren** umfassen Indikatoren wie Arbeitsbelastung, Arbeitsbedingungen, Patientenbeziehungen oder auch verschiedene Arbeitsmodelle; **4) Persönlichkeitsmerkmale** (z.B. Verträglichkeit, Offenheit); **5) Soziale Beziehungen und Wohlbefinden** wie Einsamkeitserleben oder familiäre Unterstützung; **6) Experimentelle Bedingungen** sind die aus Studie 1 (siehe 6.2.1.) bekannten experimentellen Faktoren wie KI-Autonomie, KI-Erklärbarkeit, sowie Art der medizinischen Anwendung; **7) Soziodemographische Daten** wie Alter und Geschlecht aber auch der Kontext der Praxistätigkeit (ländlich, urban).

Ergebnisse – ÄrztInnen

Die Modelle zeigten eine moderate Vorhersagekraft mit R^2 -Werten zwischen 0,317 und 0,418, wobei die Square-root LASSO-Methode die beste Leistung erbrachte. Die 5-fache Kreuzvalidierung bestätigte diese moderate Vorhersagekraft (mittleres $R^2 = 0,301$, 95% CI: 0,280-0,321) und wies auf eine relativ hohe Fehlervarianz hin (mittlerer RMSE = 0,534, 95% CI: 0,528-0,541). Die regularisierten Regressionsmodelle (LASSO, Elastic Net, Square-root LASSO) identifizierten mehrere Prädiktoren, die jedoch nicht statistisch signifikant waren. KI-Kontrolle zeigte den konsistentesten positiven Zusammenhang über alle Modelle hinweg ($\beta \approx 0,13-0,20$), gefolgt von KI-Vertrauen und KI-Nutzen. Berufliche Faktoren wie Vollzeitbeschäftigung und Arbeitsbelastung wiesen tendenziell negative, wenn auch nicht signifikante Zusammenhänge auf.

Die Random-Forest-Analyse bot eine komplementäre Perspektive: KI-Emotion war der wichtigste Prädiktor (Score 100), gefolgt von KI-Kontrolle (89,239) und KI-Vertrauen (81,745). Psychologische Faktoren (z.B. Depressionen, 62,467), berufliche Aspekte (z.B. Beschäftigungsverhältnis, 59,048) und der eingeschätzte Digitalisierungsgrad der Praxis/Einrichtung rangierten ebenfalls hoch. Zudem wurden Persönlichkeitsfaktoren (Gewissenhaftigkeit, Interesse an Neuem) und soziale Faktoren als bedeutsam identifiziert.

Die Ergebnisse deuten auf komplexe, möglicherweise nicht-lineare Beziehungen hin, was sich in den Diskrepanzen zwischen linearen Modellen und Random Forest Analyse zeigt. Die durchgehend nicht signifikanten Ergebnisse in den linearen Modellen könnten teilweise der begrenzten Stichprobengröße geschuldet sein. Trotz methodischer Einschränkungen legen vor allem die Ergebnisse der Random Forest Analyse nahe, dass KI-bezogene Einstellungen (emotionale Reaktion, wahrgenommene Kontrolle, Vertrauen) wichtige Faktoren für die Technologieakzeptanz bei ÄrztInnen sein könnten. Die mögliche Relevanz psychologischer und beruflicher Faktoren unterstreicht die Komplexität der Einflussfaktoren auf die KI-Akzeptanz im medizinischen Kontext. Es ist wichtig zu betonen, dass aufgrund der begrenzten Stichprobengröße alle diese Ergebnisse mit Vorsicht interpretiert werden müssen. Die identifizierten Muster und Zusammenhänge sollten als vorläufige Hinweise betrachtet werden, die in größeren, repräsentativeren Studien weiter untersucht und validiert werden müssen.

6.3 Projektergebnisse Handlungsempfehlungen

Ziel des Expertenworkshops war es, basierend auf den Ergebnissen der Akzeptanzstudie zum einen Implikationen für die nutzerzentrierte Entwicklung und Gestaltung von KI-Systemen, zum anderen Strategien zur Implementierung von KI-Systemen in der GKV zu erarbeiten und in konkrete Handlungsempfehlungen zu überführen. Für den Expertenworkshop, der als hybride Veranstaltung mit einer Dauer von 5 Stunden durchgeführt und zur besseren Auswertbarkeit auf Tonband aufgezeichnet wurde, konnten insgesamt 16 Personen aus den Forschungs- und Anwendungsbereichen Medizin, Entwicklung, Ethik und Soziologie gewonnen werden. Da aufgrund von parallelen Terminen nicht allen TeilnehmerInnen eine Teilnahme am gesamten Workshop möglich war, flossen lediglich die Daten der 13 dauerhaft anwesenden Personen in die Auswertung ein ($M = 42.25$ Jahre, $SD = 9.58$; 33% weiblich). Für die Datenauswertung erfolgte eine Transkription und qualitative Inhaltsanalyse in Anlehnung an Mayring (2015) mittels MAXQDA, bei der 142 Codes identifiziert wurden, die zu den vier Themenbereichen *Grundlegende Definition von KI*, *Grundlegende Rahmenbedingungen*, *Entwicklung* und *Implementierung* aggregiert und in allgemeine Anforderungen sowie Anforderungen, die spezifisch für ÄrztInnen zutreffend sind, überführt wurden. Die jeweiligen Anforderungen werden nachfolgend in Tabelle 10 stichpunktartig aufgeführt und sind ausführlich in Anhang 5 zu finden.

Tabelle 13: Handlungsempfehlungen zur Entwicklung und Implementierung von KI

Grundlegende Definitionen von KI	
<i>Allgemeine Anforderungen</i>	<i>Anforderungen für ÄrztInnen</i>
Definition von KI	KI in der Lehre/zur Schulung von medizinischem Personal
Definition der Rolle des Menschen in der KI	Kein Ersetzen von ÄrztInnen durch KI
Berücksichtigung des Einflusses von KI auf das Arbeitsfeld der Medizin	Definition des Nutzens und Einsatzbereich der KI
Wirtschaftliche Unabhängigkeit/Neutralität	(Technische) Nachvollziehbarkeit
Abgrenzung von KI zu bereits eingesetzter Tecni	KI zur Verringerung der Arbeitsbelastung
Grundlegende Rahmenbedingungen für den KI-Einsatz	
<i>Allgemeine Anforderungen</i>	<i>Anforderungen für ÄrztInnen</i>
Überwinden von bürokratischen Hürden bei der Einführung von KI	ÄrztInnen dürfen sich nicht auf KI verlassen
Berücksichtigung ethischer Anforderungen	Klärung der Haftung und Eigenbestimmung
Standardisierung innerhalb der EU	
Abhängigkeit des Vertrauens vom Nutzungskontext	
Entwicklung KI	
<i>Allgemeine Anforderungen</i>	<i>Anforderungen für ÄrztInnen</i>
Generierung der Datenbank der KI	KI als Unterstützung bei Entscheidungsfindung
Bevorzugung analytische vs. generative KI	KI als Arbeitserleichterung bei großer Datenmenge
Bottom-up-Prinzip in der Entwicklung	Einbezug bei der Entwicklung und Wissensgenerierung
Einsatz der KI in Abhängigkeit vom Anwendungsbereich	
Einfachheit der Nutzung	
Sicherstellung und Nachvollziehbarkeit der Datensicherheit	
Kompatibilität	
Standardisierung	
Berücksichtigung rechtlicher Vorgaben für die (emotionale) Gestaltung von KI	
Digitalkompetenz: Einfachheit und Verständlichkeit	
Vertrauensbildende KI-Entwicklung	
Frage nach Emotionalisierung der KI – Tendenz zu Ratio	
Berücksichtigung des Spaßfaktors der Mensch-Technik-Interaktion	
Backtracking-Funktion zur späteren Überprüfbarkeit und Vermeidung einer Blackbox-Funktion	
Einführung einer Kennzeichnung (Digital Responsibility Index)	
Implementierung KI	
<i>Allgemeine Anforderungen</i>	<i>Anforderungen für ÄrztInnen</i>
(Technische) Transparenz auf Nachfrage	Mitbestimmung bei der Implementierung
Klärung der Nachtrainierbarkeit von KI	
Berücksichtigung der Rolle der Medien bzgl. KI-Akzeptanz	
Kommunikation des Gewinns und Nutzens von KI	
Kommunikation von Informationen über KI im richtigen Ausmaß	
Durch Kommunikation und Transparenz Vertrauen schaffen	

Versicherte haben letzte Kontrollinstanz über Daten	
Versicherte haben Wahlfreiheit über Nutzung von KI in Krankenkassen	
Nutzerzentrierte Implementierung	
Kein Zwanghaftes Durchsetzen der Implementierung	
Positiver Einfluss von KI-Systemen auf Lernen und Wissensproduktion	
Berücksichtigung von Vorerfahrung/Vorkenntnissen mit KI	
Unterscheidung: Versicherte sind nicht gleich Patienten	
Höhere Akzeptanz bei Falschaussage Mensch als richtige Aussage von Maschine	
Fokussierung auf Vertrauen und Transparenz	
Berücksichtigung von Interaktionseffekten für Akzeptanz	

Darüber hinaus konnten im Rahmen der konsolidierenden Abschlussdiskussion in der Gesamtgruppe 17 zentrale Anforderungen an Entwicklung und Implementierung stichpunktartig herausgearbeitet werden, die als Priorisierung aller erarbeiteten Aspekte verstanden werden können (siehe Abbildung 14):

- Human Agency first
- Europäisch denken + handeln
- Digital & Health Literacy: Transparenz, einfache Sprache, inklusiv + intuitiv!
- KI und ÄrztInnen sollten sich gegenseitig "auf Augenhöhe" ergänzen, hinterfragen, challengen
- KI darf keine Black-Box sein: Erklärungskomponente, Nachvollziehbarkeit
- KI ist nicht KI. Um welche Methoden geht es konkret?
- Zentrale Dimensionen bei KI-Akzeptanz sollten nicht nur Nützlichkeit, Einfachheit oder Intention berücksichtigen, sondern vor allen Dingen Zuverlässigkeit/Verbindlichkeit
- Verantwortung und Entscheidungshoheit bleibt in ärztlicher Hand
- Usability
- Implementierung/Infrastruktur
- NutzerInnen-Perspektive berücksichtigen
- Aufklärung der Zielgruppen: Kompetenz + Verständnis erzeugen
- Zielgruppen Befragung zu "Was wird gebraucht?": Gap Analyse, Zielgruppe mitnehmen in der KI-Entwicklung
- Große Dynamik, viele Baustellen/ungeklärte Fragen (Recht, Ethik, Gesellschaft)
- Zentrale Stelle von Kontrolle + Vertrauen (bei UI + Medizin)
- Möglichst einheitliches deutsches/europäisches System, um Datenschutz + Kompatibilität zu garantieren
- Gute Strategien zur Implementierung nötig, sonst hilft alle Akzeptanz nichts: "Mitspracherecht" von ÄrztInnen & Versicherten

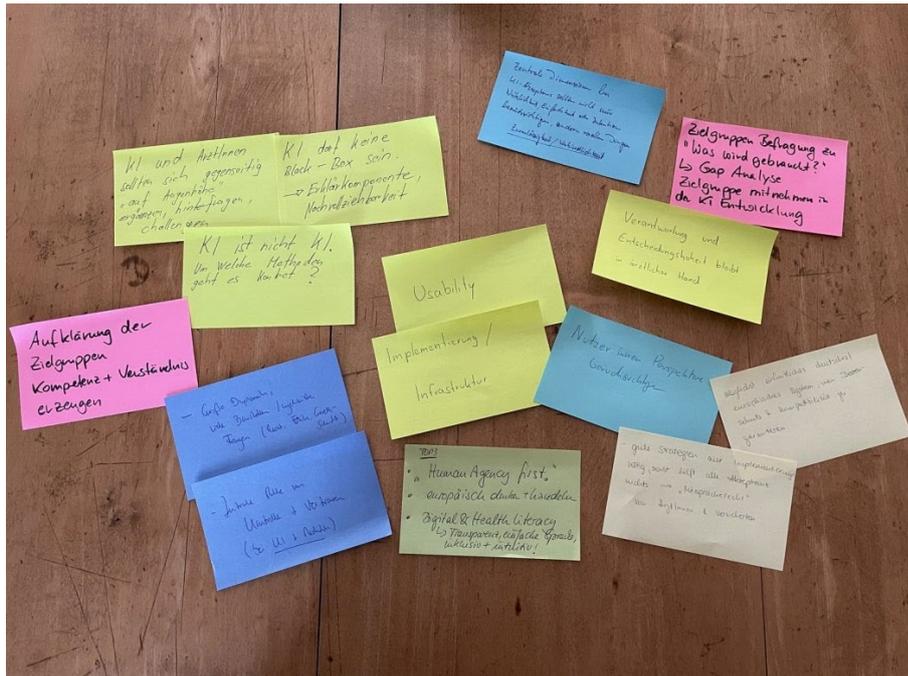


Abbildung 14: Zentrale Anforderungen im Rahmen der konsolidierenden Abschlussdiskussion

7. Diskussion der Projektergebnisse

7.1 Kategorisierung

Die Diskussion der Projektergebnisse muss unter Berücksichtigung der eingesetzten Methodik sowie der Entwicklungen im Bereich der KI-Entwicklung während der Projektlaufzeit erfolgen. So konnte aus der Literaturanalyse ein generisches und breit gefächertes Kategorisierungsinstrument entwickelt und mit Experten aus verschiedenen medizinischen und technischen Fachbereichen validiert werden, welches das strategische Management im Gesundheitswesen u.a. bei der Use Case-Entwicklung, der spezifischen Berücksichtigung ethischer Fragestellungen unterstützen und Stakeholdern im Gesundheitswesen eine Orientierung bei der Identifikation von Bereichen, die regulatorische Aufmerksamkeit fordern, bieten kann. Insbesondere der Einsatz des Kategorisierungsinstruments als Mittel zur Entwicklung von Use Cases ist jedoch stark von den adressierten Stakeholdern abhängig. So konnte das Kategorisierungssystem im Projekt zwar erfolgreich als Instrument für die systematische Ableitung von Use Cases bzw. Stimuli für die Akzeptanzstudie genutzt werden, war hier jedoch auch stark von der nicht zufallsbasierten Stichprobe und deren exemplarischer KI-Beispiele abhängig. Hier ist insbesondere vor dem Hintergrund der kleinen und nicht zufallsbasierten Stichprobe nur eingeschränkt von einer Generalisierbarkeit der ausgewählten KI-Beispiele als Stimuli auszugehen. Im Rahmen der Validierung des Kategorisierungssystems mit Experten wurde auch deutlich, dass der Einsatzzweck primär auf einer Systematisierung bereits existierender KI-Technologien und Konzepte im Sinne einer ganzheitlichen Bewertung liegt, das Kategorisierungssystem jedoch weniger für die Generierung neuer KI-Anwendungsfelder geeignet ist.

Zusammenfassend erwies sich das Kategorisierungsinstrument bereits während der Entwicklung und im Rahmen der Validierung mit Experten als dynamisches und flexibles Instrument, das um neue Technologien und Dimensionen innerhalb des Gesundheitswesens erweitert werden kann. Dieser Aspekt ist umso zentraler angesichts des schnellen Tempos technologischer Fortschritte im Bereich KI, das eine kontinuierliche Anpassung des Kategorisierungsinstruments erfordert.

7.2 Akzeptanzstudie

Die Diskussion der Projektergebnisse muss unter Berücksichtigung der Diskrepanz zwischen den angestrebten und tatsächlich erreichten Fallzahlen erfolgen. Für die Versichertenstichprobe wurde das ursprüngliche Ziel von 1.536 Personen mit 1.102 Teilnehmern zu etwa 72% erfüllt. Bei der ÄrztInnenstichprobe fällt die Abweichung deutlicher aus: Von den geplanten 475 Personen konnten nur 157 ÄrztInnen rekrutiert werden, was lediglich 33% des angestrebten Umfangs entspricht.

Diese Untererfüllung der Stichprobenziele hat Implikationen für die Validität und Generalisierbarkeit der Studienergebnisse. Die reduzierte statistische Power bedeutet, dass die Wahrscheinlichkeit, kleine Effekte in der Versichertenstichprobe zu entdecken, geringer ist als ursprünglich kalkuliert. Für die ÄrztInnenstichprobe ist die anfänglich berechnete hohe Wahrscheinlichkeit, mittlere Effekte zu identifizieren, nicht mehr gegeben. Zudem erhöht sich die Fehlerwahrscheinlichkeit in beiden Gruppen.

Die reduzierte Teilnehmerzahl bei den ÄrztInnen wirft Fragen zur Repräsentativität der Stichprobe auf. Es besteht die Möglichkeit, dass durch die Zahl nicht mehr die volle Bandbreite an Fachdisziplinen, Versorgungsumfeldern und Berufserfahrungen abgebildet wurde, was die Generalisierbarkeit der Ergebnisse einschränkt. Auch die Möglichkeiten für detaillierte Subgruppenanalysen sind durch die geringeren Fallzahlen, besonders bei den ÄrztInnen, limitiert.

Diese Einschränkungen beeinflussen sowohl die interne als auch die externe Validität der Studie. Die geringere statistische Power könnte dazu geführt haben, dass kleinere, aber möglicherweise relevante Effekte unentdeckt blieben. Insbesondere die Generalisierbarkeit der Ergebnisse auf die Gesamtpopulation der ÄrztInnen ist aufgrund der stark reduzierten Stichprobengröße kritisch zu hinterfragen.

Trotz dieser Limitationen zeigen die Studienergebnisse klare und interpretierbare Muster. Die entwickelte KI-Akzeptanz-Skala weist gute psychometrische Eigenschaften auf, und die experimentelle Studie liefert wertvolle Erkenntnisse über Einflussfaktoren auf die KI-Akzeptanz. Dennoch ist es wichtig, die Ergebnisse mit Vorsicht zu interpretieren, insbesondere wenn es um kleine Effekte oder Subgruppenanalysen geht.

7.3 Handlungsempfehlungen

Mit Blick auf die aus dem Projekt abgeleiteten Handlungsempfehlungen für den KI-Einsatz in der Gesundheitsversorgung ergeben sich aufgrund der nicht zufallsbasierten und kleinen Stichprobe ähnliche Einschränkungen hinsichtlich der Generalisierbarkeit der Ergebnisse wie hinsichtlich des Kategorisierungsinstrumentes. Nichtsdestotrotz konnten im Rahmen des Expertenworkshops nicht nur Handlungsempfehlungen, sondern konkrete Anforderungen für eine erfolgreiche Entwicklung und Implementierung von KI abgeleitet werden. Deutlich wurde hierbei insbesondere die Notwendigkeit einer Definition von KI sowie allgemeiner regulatorischer Rahmenbedingungen als Grundpfeiler für eine erfolgreiche Entwicklung und Implementierung von KI. Dieser Grundpfeiler wurde nach Durchführung des Expertenworkshops zwischenzeitlich durch den EU AI Act gelegt.

8. Verwendung der Ergebnisse nach Ende der Förderung

Die Ergebnisse des KIBA-Projekts bieten eine solide Grundlage für die weitere Entwicklung und Implementierung von KI-Anwendungen im Gesundheitswesen. Nach Abschluss der Förderung ergeben sich mehrere vielversprechende Ansatzpunkte zur Fortführung und Weiterentwicklung der Forschung sowie zur praktischen Anwendung der gewonnenen Erkenntnisse.

Ein zentrales Ergebnis ist die Entwicklung einer multidimensionalen Skala zur Messung der individuellen KI-Akzeptanz im Gesundheitskontext. Diese Skala, die die Faktoren Vertrauen, Angst, Kontrolle und wahrgenommenen Nutzen umfasst, sollte weiter validiert und für verschiedene Anwendungsbereiche im Gesundheitswesen optimiert werden. Ziel ist es, ein standardisiertes und robustes Instrument zur Verfügung zu stellen, das sowohl in der Forschung als auch in der Praxis eingesetzt werden kann, um die Akzeptanz von KI-Anwendungen bei verschiedenen Stakeholdern zuverlässig zu messen.

Die identifizierten Einflussfaktoren auf die KI-Akzeptanz bieten wichtige Ansatzpunkte für die Entwicklung zielgruppenspezifischer Schulungs- und Informationsprogramme. Für ÄrztInnen scheinen dabei insbesondere die Arbeitsbedingungen und die Art der Beschäftigung relevant zu sein, während bei Versicherten Faktoren wie Alter und allgemeine Technikmotivation eine größere Rolle spielen. Diese Erkenntnisse sollten genutzt werden, um maßgeschneiderte Programme zu entwickeln, die das Verständnis und die Akzeptanz von KI im Gesundheitswesen fördern.

Bei der Implementierung von KI-Systemen in der Gesundheitsversorgung sollten die Erkenntnisse zu Präferenzen bezüglich KI-Autonomie und Erklärbarkeit berücksichtigt werden. Die Studie zeigte eine höhere Akzeptanz für Systeme, bei denen Menschen die finalen Entscheidungen treffen, im Vergleich zu vollautonomen KI-Entscheidungen. Dies unterstreicht die Bedeutung eines "Human-in-the-Loop"-Ansatzes bei der Gestaltung von KI-Anwendungen im Gesundheitswesen.

Die beobachteten Unterschiede in der KI-Akzeptanz zwischen ÄrztInnen und Versicherten weisen auf die Notwendigkeit weiterer, vertiefender Forschung hin. Zukünftige Studien sollten diese Unterschiede genauer untersuchen und Strategien zur Überbrückung möglicher Akzeptanzlücken zwischen verschiedenen Stakeholdergruppen entwickeln.

Um den langfristigen Beitrag von KI zur Weiterentwicklung der GKV-Versorgung zu evaluieren, wären Langzeitstudien wertvoll. Diese könnten die tatsächlichen Auswirkungen der KI-Implementierung auf die Versorgungsqualität, Arbeitsbedingungen und Patientenzufriedenheit untersuchen und somit eine evidenzbasierte Grundlage für zukünftige gesundheitspolitische Entscheidungen liefern.

Für die Überführung der Projektergebnisse in die Versorgung ist die Entwicklung von Best Practices für die Einführung und Kommunikation von KI-Anwendungen im Gesundheitswesen ein wichtiger nächster Schritt. Dabei sollten die Erkenntnisse zur Akzeptanzförderung, wie die Bedeutung von Transparenz und menschlicher Kontrolle, berücksichtigt werden.

Abschließend ist zu betonen, dass die Methodik und Erkenntnisse des KIBA-Projekts das Potenzial haben, auf andere Bereiche des Gesundheitswesens oder andere innovative Technologien übertragen zu werden. Dies könnte dazu beitragen, einen umfassenderen Ansatz zur Förderung der Akzeptanz und effektiven Nutzung neuer Technologien in der Gesundheitsversorgung zu entwickeln.

Der nächste konkrete Schritt zur Überführung der Projektergebnisse in die Regelversorgung wäre die Entwicklung von Pilotprojekten in Zusammenarbeit mit Krankenkassen, Ärzteverbänden und Patientenorganisationen. Diese Pilotprojekte könnten die entwickelten Akzeptanzförderungsstrategien in der Praxis testen und evaluieren, um einen evidenzbasierten Weg für die breitere Implementierung von KI in der Gesundheitsversorgung zu ebnen.

9. Erfolge bzw. geplante Veröffentlichungen

- Erste Ergebnisse aus dem Projekt wurden mit zwei Postern auf der Jahrestagung der DGGG (Deutsche Gesellschaft für Gerontologie und Geriatrie) in Vechta im September 2023 aufbereitet und vorgestellt.
- Vortrag **KI in der Altersforschung: Zwischen Hype und Realität – Aktuelle Fortschritte, Herausforderungen und Chancen**, Kongress der Deutschen Gesellschaft für Gerontologie und Geriatrie (11.-14.09.24), Symposium zu Gesundheit und Gesundheitsversorgung, Ort: Universität Kassel
- Vortrag **Digitalisierung und Alter(n): Verloren im KI-Dschungel?**, Journée de Gérontologie 2024 (02.10.2024), Association luxembourgeoise de Géronto-Gériatrie, Ort: Dommeldage, Luxemburg
- Conference Paper **Strategizing AI in Healthcare: A Multi-dimensional Blueprint for Transformative Decision-Making in Clinical Settings** (Martina Simon, Stefan Kamin, Andreas Hamper, Thomas Wittenberg, Stephanie Schmitt-Rüth), 8th Annual Conference of the German Society for Biomedical Engineering (18.-20.09.24)
- **Exploring the Role of Openness to new experiences and Subjective Technology Adaptivity in Digital Technology Use Among Older Adults** (Stefan T. Kamin, Nora M. Degen, Anja Beyer, Martina Simon, Stephanie Schmitt-Rüth, Frieder R. Lang)
- Buch-Kapitel „Acceptance of AI by clinical experts“ im geplanten 4. Band der der Health Academy »KI, Medizintechnik und Ethik«
- **Acceptance of Artificial Intelligence in Healthcare** (Stefan T. Kamin, Helena S. Hösch, Nora M. Degen, Frieder R. Lang, in Vorbereitung)

10. Literaturverzeichnis

Auer, C., Hollenstein, N., & Reumann, M. (2019). Künstliche Intelligenz im Gesundheitswesen. In *Gesundheit digital* (pp. 33-46).

Bundesregierung. (2018). *Strategie Künstliche Intelligenz der Bundesregierung*.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of technology. *MIS Quarterly*, 13, 319–340.

Hackel, M., & Klebl, M. (2008). Qualitative Methodentriangulation bei der arbeitswissenschaftlichen Exploration von Tätigkeitssystemen. *Forum Qualitative Sozialforschung*, 9(3).

Kamin, S. T., Lang, F. R., & Beyer, A. (2017). Subjective Technology Adaptivity Predicts Technology Use in Old Age. *Gerontology*, 63, 385–392. doi:10.1159/000471802

Kett, S., Ringler, A., Kamin, S. T., Lang, F. R., & Wartzack, S. (2015). Kritische Würdigung eines Werkzeugs zur Messung von Nutzerexklusion im Produktdesign. In D. Krause (Ed.), *Design for X* (neue Ausg ed., pp. 13–25). Hamburg: TuTech Innovation.

Lucke, D. (1995). *Akzeptanz: Legitimität in der „Abstimmungsgesellschaft“*. VS Verlag für Sozialwissenschaften. ISBN: 978-3-8100-1496-2.

Manzeschke, A., Weber, K., Rother, E., & Fangerau, H. (2013). *Ethische Fragen im Bereich Altersgerechter Assistenzsysteme*.

Mayring, P. (2002). *Einführung in die qualitative Sozialforschung*. Weinheim: Beltz.

Neyer, F. J., Felber, J., & Gebhardt, C. (2016). Kurzsкала Technikbereitschaft (TB, technology commitment): ZIS–GESIS Leibniz Institute for the Social Sciences.

Akronym: KI-BA

Förderkennzeichen: 01VSF20016

Seifert, A., Kamin, S. T., & Lang, F. R. (2020). Technology Adaptivity Mediates the Effect of Technology Biography on Internet Use Variability. *Innovation in Aging*, 4(2). doi:10.1093/geroni/igz054

Stöber, C., Williger, B., Meerkamm, H., & Lang, F. R. (Eds.). (2012). *Leitfaden für die altersgerechte Produktentwicklung*. Stuttgart: Fraunhofer Verlag.

Tausch, A., & Menold, N. (2015). *Methodische Aspekte der Durchführung von Fokusgruppen in der Gesundheitsforschung*. 51.

11. Anhang

Anhang 1: Ergebnisdokumentation AP 1 – Kategorisierung

Anhang 2: AP 1 – Konzeptmatrizen und verwendete Literatur

Anhang 3: AP 2 – Stimulusdesign Texte

Anhang 4: AP 3 – Studiendesign

Anhang 5: Ergebnisdokumentation AP 9 - Handlungsempfehlungen

12. Anlagen

Anlage 1: AP 3 – Codebuch Akzeptanzstudie ÄrztInnen

Anlage 2: AP 3 – Codebuch Akzeptanzstudie Versicherte

Ergebnisdokumentation

01VSF20016 - KI-BA

Künstliche Intelligenz in der Versorgung-
Bedingung der Akzeptanz von
Versicherten und ÄrztInnen

AP 1 „Kategorisierung von KI in der Versorgung“

Inhalt

II Abbildungsverzeichnis.....	3
III Tabellenverzeichnis.....	3
1. Zusammenfassung.....	4
2. AP 1 – Projektplan.....	4
3. Literature Review zur Vorbereitung einer Kategorisierung von KI in der Versorgung.....	5
3.1 Methodik.....	5
3.2 Ergebnis.....	10
4. Experteninterviews zur Validierung der Kategorisierung von KI in der Versorgung.....	17
4.1 Methodik.....	17
4.2 Ergebnis.....	18
4.2.1 Demografika.....	18
4.2.2 Definitionen von KI.....	19
4.2.3 Erste Eindrücke zum Kategorisierungssystem.....	19
4.2.4 Kategorisierung eigener Beispiele im Kategorisierungssystem.....	20
4.2.5 Erklärung eigener KI-Beispiele für KI-Neulinge.....	25
4.2.6 Evaluation des Kategorisierungssystems hinsichtlich Vollständigkeit, Verständlichkeit und Interdisziplinarität.....	26
4.2.7 Finales Kategorisierungssystem.....	28
5. Online-Fragebogen zur Stimulusauswahl auf Basis des Kategorisierungsinstruments....	29
5.1 Methodik.....	29
5.2 Ergebnis.....	29
5.2.1 Bewertung der KI-Beispiele hinsichtlich Verständlichkeit, Relevanz und Entlastungsfunktion.....	30
5.2.1.1 Bewertung der KI-Beispiele – Prävention.....	30
5.2.1.2 Bewertung der KI-Beispiele – Diagnostik.....	32
5.2.1.3 Bewertung der KI-Beispiele – Therapie.....	34
5.2.2 Auswahl und Priorisierung von KI-Beispielen für KI-Akzeptanzstudie.....	37
5.2.2.1 Auswahl und Priorisierung von KI-Beispielen – Prävention.....	37
5.2.2.2 Auswahl und Priorisierung von KI-Beispielen – Diagnostik.....	39
5.2.2.3 Auswahl und Priorisierung von KI-Beispielen – Therapie.....	41
5.2.3 Auswahl des Nutzungsszenarios aus Arzt- bzw. Patientensicht.....	43
Literatur.....	48
Anhang.....	50

II Abbildungsverzeichnis

Abbildung 1 Literatursuche und -auswahlprozess (eigene Darstellung in Anlehnung an Schryen (2015), Wolfswinkel, Furtmueller und Wilderom (2013), Zhang Babar und Tell (2011) und vom Brocke et. al. (2009)	13
Abbildung 2 Überblick über Dubletten und Relevanzabschätzung	14
Abbildung 3 1. Entwurf eines Kategorisierungsinstruments für KI in der Versorgung	16
Abbildung 4 Kategorisierungssystem als Grundlage für Experteninterviews.....	18
Abbildung 5. Übersicht aller KI-Beispiele und Einordnung ins Kategorisierungssystem.	21
Abbildung 6. Ergebnisse für die ausgewählten KI-Beispiele - Bereich Prävention	43
Abbildung 7. Ergebnisse für die ausgewählten KI-Beispiele - Bereich Diagnostik.....	44
Abbildung 8. Ergebnisse für die ausgewählten KI-Beispiele - Bereich Therapie.....	45
Abbildung 9. Ableitung der Auswahl des KI-Beispiels - Bereich Prävention	46
Abbildung 10. Ableitung der Auswahl des KI-Beispiels - Bereich Diagnostik.....	46
Abbildung 11. Ableitung der Auswahl des KI-Beispiels - Bereich Therapie	47

III Tabellenverzeichnis

Tabelle 1 Klassifizierung der durchgeführten Literaturanalyse zur Kategorisierung von KI in der Versorgung (eigene Darstellung in Anlehnung an vom Brocke et al. (2015), Fettke (2006), Cooper und Hedges (2009) und Cooper (1988))	7
Tabelle 2 Erklärungen zu verwendeten Merkmalen	8
Tabelle 3 Überblick über systematische Literaturanalysen zum Thema KI im Gesundheitswesen	11
Tabelle 4 Übersicht über die verwendeten Literaturdatenbanken.....	12
Tabelle 5 Überblick über Suchzeitpunkt und Suchfilter	12
Tabelle 6 Zusätzliche Suchstrategien und Literaturquellen	14
Tabelle 7 Ein- und Ausschlusskriterien für die Literatursuche	15
Tabelle 8 Soziodemografische Angaben.....	18
Tabelle 9. Einschätzung des Kategorisierungssystems.....	26
Tabelle 10. Finales Kategorisierungssystem.....	28
Tabelle 11. Einschätzung der KI-Beispiele aus der Prävention bzgl. Verständlichkeit	30
Tabelle 12. Einschätzung der KI-Beispiele aus der Prävention bzgl. Relevanz.....	30
Tabelle 13. Einschätzung der KI-Beispiele aus der Prävention bzgl. Entlastungsfunktion	31
Tabelle 14. Einschätzung der KI-Beispiele aus der Diagnostik bzgl. Verständlichkeit.....	32
Tabelle 15. Einschätzung der KI-Beispiele aus der Diagnostik bzgl. Relevanz.....	32
Tabelle 16. Einschätzung der KI-Beispiele aus der Diagnostik bzgl. Entlastungsfunktion.....	33
Tabelle 17. Einschätzung der KI-Beispiele aus der Therapie bzgl. Verständlichkeit	34
Tabelle 18. Einschätzung der KI-Beispiele aus der Therapie bzgl. Relevanz	35
Tabelle 19. Einschätzung der KI-Beispiele aus der Therapie bzgl. Entlastungsfunktion	36
Tabelle 20. Auswahl der am besten geeigneten KI-Beispiele aus der Prävention	37
Tabelle 21. Auswahl der am besten geeigneten KI-Beispiele aus der Diagnostik.....	39
Tabelle 22. Auswahl der am besten geeigneten KI-Beispiele aus der Therapie	41

1. Zusammenfassung

Das nachfolgende Dokument liefert einen Überblick über die Aktivitäten und Arbeitsergebnisse von AP 1 (Kategorisierung von KI in der Versorgung). Ausgehend von einer Darstellung der im Projektplan definierten Inhalte und zu erreichenden Ergebnisse (Kapitel 2) werden in Kapitel 3 Vorgehen und Ergebnisse des systematischen Reviews der vorhandenen Literatur zur Kategorisierung von KI-Anwendungen in der medizinischen Versorgung beschrieben. Kapitel 4 stellt die Vorgehensweise und Ergebnisse der Experteninterviews sowie das daraus abgeleitete Kategorisierungsinstrument für KI in der medizinischen Versorgung dar. Kapitel 5 liefert abschließend einen Überblick über die Methodik und die Ergebnisse des Online-Fragebogens, der zur Ableitung von Stimuli für die Akzeptanzstudie auf Basis des Kategorisierungsinstruments entwickelt und eingesetzt wurde.

2. AP 1 – Projektplan

Inhaltsbeschreibung: Die Anwendungen von KI für den Gesundheitsbereich sind mannigfaltig und können im Hinblick auf mehrere Dimensionen kategorisiert werden. Je nach Betrachtungswinkel kann die Kategorisierung z.B. nach Fachdisziplin (z.B. Radiologie, Endoskopie, Dermatologie, Geriatrie, Allgemeinmedizin), Einsatz in der Versorgung (z.B. Prävention, Diagnose, Therapie, Nachsorge, oder Pflege), KI-Forschungsfeld (z.B. Automatisches oder adaptives Lernen, Explainable AI, Semantik oder mathematische Optimierung) oder Zielgruppe (z.B. Facharzt, Allgemeinarzt, Fachkraft, Patient, Verwaltung, Angehörige) unterschieden werden. Für die Bildung eines systematischen Überblicks wird ein systematisches Review der vorhandenen Literatur zur Kategorisierung von KI-Anwendungen in der medizinischen Versorgung durchgeführt. Die aus der Literaturanalyse abgeleiteten Kategorisierungsansätze werden im Rahmen von Experteninterviews mit Experten aus der Gesundheitswirtschaft, Medizin, KI-Forschung, Medizintechnik und Versorgungsforschung diskutiert, eingegrenzt und erweitert. Zusammen mit den Experten werden exemplarische Beispiele für Vertreter der jeweiligen Kategorien erhoben und beschrieben. Ziel der Entwicklung ist ein mehrdimensionales Kategorisierungsinstrument, dass der interdisziplinären Sicht (Sicht von Patienten, Ärzten, Fachdisziplinen, Versorgungssektoren, etc.) gerecht wird und diese Sichtweisen als Dimensionen abbildet

Ergebnisse:

- Kategorisierungsinstrument für KI in der Versorgung
- Mehrdimensionaler Überblick über das Spektrum KI-basierter Anwendungen im Gesundheitswesen
- Beschreibung exemplarischer Beispiele für Vertreter der Kategorien

Meilensteine:

- Meilenstein 3: Vorbereitung Kategorisierung KI in der Versorgung
- Meilenstein 4: Kategoriensystem erstellt

3. Literature Review zur Vorbereitung einer Kategorisierung von KI in der Versorgung

3.1 Methodik

Bei einer strukturierten Literaturübersicht handelt es sich um eine systematische, explizite und reproduzierbare Methode zur Identifizierung, Interpretation, Bewertung und Zusammenfassung der vorhandenen Literatur von Forschern, Wissenschaftlern und Praktikern zu einem bestimmten Themenbereich. Auf diese Weise tragen Literaturübersichten dazu bei, Forschungsdefizite zu identifizieren und offene Forschungsfragen abzuleiten. Des Weiteren dienen Literaturübersichten dazu, ein Gerüst und eine solide Forschungsgrundlage zu schaffen, indem sie die erneute Erforschung des bestehenden Kenntnisstandes verhindern und eine effektive Wiederverwendung des vorhandenen Wissens gewährleisten. Darüber hinaus tragen sie dazu bei, bereits aufgestellte Theorien zu verifizieren und neue Theorien zu entwickeln, indem bestehende Theorien adaptiert, neue Theorien aufgestellt oder verschiedene bestehende Theorien zusammengeführt werden.¹

Zur Erreichung dieser Ansprüche haben sich in der Wissenschaft die von Harris M. Cooper entwickelten nachfolgenden fünf Phasen zur systematischen Durchführung einer Literaturliteraturanalyse als standardisiertes Ablaufschema bewährt:

1. **Problemformulierung:** Gegenstand der ersten Phase einer Literaturliteraturanalyse ist das Aufstellen der zu überprüfenden Forschungsfrage und der dazugehörigen Teilfragestellungen.
2. **Literaturliteratursuche:** Zur Beantwortung der Forschungsfrage soll in der zweiten Phase der Literaturliteraturanalyse die entsprechende Literatur aus relevanten Quellen gesucht werden.
3. **Literaturliteraturauswahl:** Im Rahmen der dritten Phase erfolgt eine Relevanzprüfung sowie eine Qualitätsbewertung der selektierten Literatur.
4. **Analyse und Interpretation:** Die vierte Phase der Literaturliteraturanalyse ist gekennzeichnet durch die Synthese, das Strukturieren und Kategorisieren der relevanten Literaturliteraturergebnisse. Darüber hinaus sind eventuelle Forschungslücken aufzudecken, damit verbundene ungeklärte Forschungsfragen zu ermitteln und ein abschließendes Fazit zu erarbeiten.
5. **Präsentation der Ergebnisse:** Abschließend werden in der letzten Phase die gewonnenen Erkenntnisse mit Hilfe von Tabellen und Abbildungen in angemessener Form aufbereitet, beschrieben und visualisiert.²

Literaturliteratursuche und Literaturliteraturauswahl:

Die Literaturliteraturanalyse zur Beantwortung der in der Problemstellung definierten Forschungsfragen folgte vor diesem Hintergrund einem iterativen Prozess der Literaturliteraturrecherche und -auswertung, der die Menge an verfügbarer Literatur auf eine überschaubare Menge unter Wahrung eines möglichst **hohen Abdeckungsgrades** reduzieren sollte. Ebenso verhält es sich mit der **Art der zu berücksichtigten Literatur**, zu der sowohl wissenschaftliche als auch nicht-wissenschaftliche Literatur gehört, wie auch mit der Auswahl der **Literaturliteraturdatenbanken**, mit der

¹ Vgl. Boell und Cecez-Kecmanovic (2015), Fink (2014), Levy und Ellis (2006), Schryen (2015), Schryen et al. (2017), vom Brocke et. al. (2009), Webster und Watson (2002)

² Vgl. Boell und Cecez-Kecmanovic (2015), Cooper (1982), Cooper und Hedges (2009), Fettke (2006), Schryen (2015), vom Brocke et al. (2009), Wolfswinkel, Furtmueller und Wilderom (2013).

insbesondere die Erfassung qualitativer Literatur sichergestellt wird, sowie mit den **sonstigen Literaturquellen**, die ergänzend herangezogen wurden.³

Der ausgewählte Ablauf der Literatursuche und -auswertung wurde in Abschnitte unterteilt: Schlüsselwortsuche, Relevanzeinschätzung, Auswertung und zusätzliche Suchen / Verfeinerung der Suche. Die Recherche beginnt mit der Formulierung passender **Schlüsselwörter**, anhand derer die vorab definierten Forschungsfragen untersucht werden.⁴ In Abhängigkeit der zu berücksichtigenden Literaturart sind dann relevante Literaturquellen zu identifizieren. Von zentraler Relevanz sind hierbei **Literaturdatenbanken**, die bei einer Literaturanalyse zu berücksichtigen sind. Bei der Suche sollten auch **sonstige Literaturquellen** herangezogen werden (z.B. Expertenbefragungen, Forschungsregistern oder im Internet allgemein).⁵ Abschließend erfolgt eine **Schlüsselwortsuche** in sämtlichen Literaturquellen mithilfe zuvor definierter Suchbegriffe.⁶ Eine möglichst weitreichende Abdeckung ist ausschlaggebend für die Güte einer Literaturanalyse, weshalb verschiedene Literaturdatenbanken durchsucht werden, wobei zu beachten ist, dass diese individuell formulierte Suchanfragen mit denselben Stichworten erfordern.⁷ Die **Suchergebnisse** der verschiedenen Datenbanken sind im Anschluss konsistent **zu exportieren und in einheitlichem Format zusammenzuführen**. Nach der Suche werden im Rahmen der **Relevanzeinschätzung** irrelevante Artikel aussortiert: Dabei werden **Duplikate** entfernt, da verschiedene Literaturdatenbanken teilweise gleiche Quellen erfassen. Schließlich werden die Artikel zunächst auf Basis ihres **Titels**, dann auf Basis ihres **Abstracts** und letztlich auf Basis ihres **Volltextes** hinsichtlich ihrer Relevanz für die Forschungsfragen hin selektiert.⁸ Zum **Abruf der Volltexte** wurde im Rahmen der Literatursuche ein lizenzierter Zugang der Universitätsbibliothek Friedrich-Alexander-Universität Erlangen-Nürnberg verwendet, wodurch ein erheblicher Teil der gewünschten Literatur bezogen werden konnte.⁹ Artikel, die nicht direkt beschafft werden konnten, wurden ausgeschlossen. Nachdem die Wahrscheinlichkeit gering ist, auf Anhieb geeignete Suchbegriffe zu ermitteln, welche die zu untersuchende Forschungsfrage umfassend abbilden, ist es unter Umständen notwendig, die Suchbegriffe mit fortschreitender Literaturlauswertung mehrfach an die mittlerweile gesammelten Erkenntnisse zu adaptieren. Als weiterer Schritt der Literaturanalyse dient hierfür die **Verfeinerung der Suche**.¹⁰ Um einen möglichst breiten Erfassungsgrad zu erzielen, werden darüber hinaus **weitere Sucharten** (z.B. Rückwärtssuche, Vorwärtssuche, Autorensuche) verfolgt, die jeweils bzgl. ihrer Relevanz zu bewerten sind und ggf. zu einer iterativen Anpassung der Schlüsselwortsuche führen können.¹¹

In Bezug auf die **Form der zu berücksichtigenden Literatur** erfolgte die Erhebung des aktuellen Forschungsstandes im Rahmen der durchgeführten Literaturanalyse folgendermaßen: Gemäß der Anforderung, insbesondere qualitativ hochwertige wissenschaftliche Literatur zu berücksichtigen, wird empfohlen, in erster Linie von Experten begutachtete (*peer-reviewed*) Literatur zu berücksichtigen, die in **wissenschaftlichen Zeitschriften** und **Tagungsbänden wissenschaftlicher Konferenzen** veröffentlicht wurde.¹² Ferner wurde wissenschaftliche Literatur anderer

³ Vgl. vom Brocke et al. (2015) und vom Brocke et al. (2009).

⁴ Vgl. vom Brocke et al. (2015).

⁵ Vgl. Schryen (2015) und vom Brocke et al. (2015).

⁶ Vgl. Levy und Ellis (2006).

⁷ Vgl. vom Brocke et al. (2009), vom Brocke et al. (2015), Levy und Ellis (2006), Brereton et al. (2007).

⁸ Vgl. Wolfswinkel, Furtmueller und Wilderom (2013) und vom Brocke et al. (2009)

⁹ Vgl. Schryen (2015).

¹⁰ Vgl. Boell und Cecez-Kecmanovic (2015).

¹¹ Vgl. Levy und Ellis (2006) und Webster und Watson (2002).

¹² Vgl. Schryen (2015), vom Brocke et al. (2009), Kitchenham und Charters (2007), Levy und Ellis (2006) und Webster und Watson (2002).

Publikationsformen wie **Monografien**, **Sammelwerken** oder **Berichten** berücksichtigt. Ergänzend zur Empfehlung wurden dementsprechend auch **nicht-wissenschaftliche Publikationen** (Marktstudien, Whitepapers, Firmenpublikationen, nicht-wissenschaftliche Fachzeitschriften, Tageszeitung etc.)/ **graue Literatur** berücksichtigt, um den Stand der Praxis umfassend abzubilden und einem möglichen Publikationsbias durch alleinige Auswertung wissenschaftlicher Artikel aus hochrangigen Journals entgegenzuwirken.¹³

Auswahl der Literaturdatenbanken

Um Art, Ausmaß und Güte zu beurteilen, bietet sich die Verwendung einer Taxonomie zur Klassifizierung von Literaturanalysen an, erstmals von Harris M. Cooper vorgeschlagen und später in Form eines morphologischen Kastens dargestellt.¹⁴ Auf Basis dieses Rahmenwerks wurde die in Tabelle 1 dargestellte Taxonomie erstellt, um die durchgeführte Literaturanalyse klassifizieren und bewerten zu können.

Tabelle 1 Klassifizierung der durchgeführten Literaturanalyse zur Kategorisierung von KI in der Versorgung (eigene Darstellung in Anlehnung an vom Brocke et al. (2015), Fettke (2006), Cooper und Hedges (2009) und Cooper (1988))

Fokus	Forschungsergebnisse	Forschungsmethoden	Theorien / Konzepte		Praktiken / Anwendungen
Ziele	Integration und Synthese der relevanten Literatur		Kritische Analyse		Identifikation zentraler Herausforderungen
Perspektive	neutral / objektiv		parteiisch / subjektiv		
Abdeckung	vollständig	vollständig mit selektiver Zitation	repräsentativ		Schlüsselartikel
Suchstrategie	Schlüsselwortsuche	Suchaufträge	Rückwärtsuche	Vorwärtsuche	Autorensuche
Literaturarten	Wissenschaftlich		Nicht-wissenschaftlich		Graue Literatur
Wissenschaftliche Literatur	Zeitschriften	Konferenzen	Monographien	Sammelwerke	Forschungsberichte / Policy Paper
Nichtwissenschaftliche Literatur	Marktstudien	Standards / Normen	Whitepaper / Firmenschriften		Sonstige Zeitschriften
Datenbanken	Volltextdatenbanken		Bibliographische Datenbanken		Zitationsdatenbanken
Sonstige Quellen	Experten	Kollegen	Forschungsregister		Internet
Gliederung	historisch		thematisch / konzeptuell		methodisch

¹³ Vgl. Schryen (2015), Onwuegbuzie, Leech und Collins (2012) und Petticrew und Roberts (2006), Rothstein und Hopewell (2009)

¹⁴ Vgl. Cooper und Hedges (2009), Cooper (1988), vom Brocke et al. (2015), vom Brocke et al. (2009) und Fettke (2006).

Zielgruppe	Fachleute	Wissenschaft- ler	Praktiker / Ent- scheider	Allg. Öffentlich- keit
------------	-----------	----------------------	------------------------------	---------------------------

Ergänzend erklärt Tabelle 2 die verwendeten Eigenschaften und Kategorien.

Tabelle 2 Erklärungen zu verwendeten Merkmalen

Merkmal	Ausprägung	Im Rahmen der Literaturanalyse werden ...
Fokus	Forschungsergebnisse	... Artefakte, Theorien oder empirische Befunde berücksichtigt.
	Praktiken & Anwendungen	... gesammelte praktische Erfahrungen und Anwendungsbereiche berücksichtigt, in denen die analysierte Literatur anzusiedeln ist.
Ziele	Integration und Synthese der relevanten Literatur	... Zusammenfassungen, Strukturierungen und Kategorisierungen der Ergebnisse der relevanten Literatur berücksichtigt, wobei in der Regel verschiedene Arbeiten zueinander in Beziehung gesetzt werden.
Perspektive	neutral / objektiv	... keine bestimmten theoretische Positionen bezogen, sondern bei Analyse der relevanten Literatur eine neutrale Stellung eingenommen.
Abdeckung	repräsentativ	... Quellen aufgrund definierter Merkmale in unterschiedliche Gruppen eingeteilt, aus welchen jeweils repräsentative Arbeiten frei ausgewählt und beschrieben werden.
Suchstrategie	Schlüsselwortsuche	... Suchen auf Basis von definierten Suchbegriffen in ausgewählten Literaturdatenbanken durchgeführt (i.d.R. in mehreren Literaturdatenbanken, um einen hohen Abdeckungsgrad zu erreichen).
	Rückwärtssuche	... zur Identifikation von weiteren relevanten Artikeln das Literaturverzeichnis von bereits als relevant eingestuften Artikeln ausgewertet.
	Autorensuche	... zur Identifikation von weiteren relevanten Artikeln spezielle Zitationsdatenbanken nach Autoren durchsucht, welche bereits als relevant eingestufte Artikel verfasst haben (vorwärts- oder rückwärtsgerichtet).
Literaturarten	Wissenschaftlich	... wissenschaftliche Quellen berücksichtigt, die in unterschiedlichen Publikationsformen vorliegen kann, zur Erfassung des Stands der Wissenschaft.
	Nichtwissenschaftlich	... nichtwissenschaftliche Quellen berücksichtigt, die in unterschiedlichen Publikationsformen vorliegen kann, zur Erfassung des Stands der Praxis.
	Graue Literatur	... wissenschaftliche und nichtwissenschaftliche Quellen berücksichtigt, die in unterschiedlichen Publikationsformen vorliegen können, sofern diese nicht über das klassische, kommerzielle Verlagswesen oder den Buchhandel bezogen werden kann.

Nicht-/Wissenschaftliche Literatur	Zeitschriften	... wissenschaftliche Artikel berücksichtigt, die in wissenschaftlichen Zeitschriften veröffentlicht wurden und deren Qualität in der Regel durch ein Begutachtungsverfahren überprüft wurde.
	Konferenzen	... wissenschaftliche Artikel berücksichtigt, die in Konferenzbänden von wissenschaftlichen Konferenzen veröffentlicht wurden und deren Qualität meist mittels Begutachtungsverfahren überprüft wurde.
	Monographien	... umfassende, selbständige wissenschaftliche Werken wie bspw. Habilitations- und Dissertationsschriften berücksichtigt.
	Sammelwerke	... wissenschaftliche Artikel berücksichtigt, die in Sammelwerken veröffentlicht wurden.
	Forschungsberichte / Policy Paper	... wissenschaftliche Artikel berücksichtigt, die oft kein Begutachtungsverfahren durchlaufen haben und als graue Literatur veröffentlicht wurden.
	Whitepaper / Firmenschriften	... Whitepapers und sonstigen Firmenschriften berücksichtigt, die von Unternehmen veröffentlicht wurden.
Datenbanken	Volltextdatenbanken	... digitale Literaturdatenbanken berücksichtigt, welche Literatur im Volltext durchsuchbar und zugänglich machen.
	Bibliographische Datenbanken	... digitale Literaturdatenbanken berücksichtigt, welche Literatur auf Basis bibliographische Angaben wie dem Titel und der Zusammenfassung eines Artikels durchsuchbar macht.
	Zitationsdatenbanken	... digitale Literaturdatenbanken berücksichtigt, welche Zitationsanalysen ermöglichen, indem diese die Zitationsbeziehungen zwischen unterschiedlichen Publikationen erfassen und durchsuchbar machen.
Sonstige Quellen	Internet	... Internetsuchmaschinen zur Identifikation zusätzlicher relevanter Literatur über die durchsuchten Literaturdatenbanken hinaus berücksichtigt.
Gliederung	thematisch / konzeptuell	... relevante Arbeiten, die sich auf dieselben abstrakten Konzepte bezieht, vergleichend und zusammengehörig besprochen.
Zielgruppe	Fachleute	... Spezialisten einer ausgewählten Teildisziplin angesprochen.
	Wissenschaftler	... alle Forscher einer wissenschaftlichen Disziplin angesprochen.

Die Literaturanalyse hat den Anspruch, die identifizierten Konzepte, Kategorien etc. in geeigneter Form mittels Tabellen und Abbildungen zu beschreiben, darzustellen und zu visualisieren. Webster und Watson (2002) raten in diesem Kontext, eine systematische Literaturanalyse nach einem konzeptzentrierten Ansatz zu strukturieren und dabei eine **Konzeptmatrix** zur Ergebnisdarstellung zu verwenden, mittels derer die identifizierten Konzepte der relevanten Literatur in geeigneter Weise zusammengefasst werden.¹⁵

¹⁵ Webster und Watson (2002)

3.2 Ergebnis

3.2.1 Problemformulierung

Zur Erhebung des Stands der Technik bei KI-Anwendungen für das Gesundheitswesen im Allgemeinen wurde die folgende Forschungsfrage formuliert, die durch die systematische Literaturübersicht beantwortet werden konnte:

RQ 1: Welchen Beitrag leistet die Literatur für die Entwicklung eines Kategorisierungsinstrumentes zur Definition von Dimensionen und Kategorien für KI-Anwendungen im Gesundheitswesen?

Zur Schärfung und Präzisierung der Frage wurden schrittweise Antworten auf folgende Teilfragen gesucht:

- *RQ 1.1: Welche Kategorisierungen, Systematisierungen oder Taxonomien bestehen bereits in der Literatur des Forschungsfelds KI im Gesundheitswesen?*
- *RQ 1.2: Welche Kategorisierungsansätze in Bezug auf KI in der Versorgung lassen sich aus der Literatur ableiten?*
- *RQ 1.3: Welche Ausprägungen werden als Auswahlobjekte herangezogen und in welche Kategorien lassen sich diese einordnen?*

3.2.2 Literatursuche und Literaturauswahl

Vor der Durchführung der systematischen Literaturanalyse zur Kategorisierung von KI-Anwendungen wurde geprüft, ob bereits entsprechende Übersichtsarbeiten existieren, auf die sich die vorliegende Literaturrecherche stützen kann. Vor diesem Hintergrund sind in Tabelle 3 insgesamt drei systematische Literaturanalysen zum Thema Künstliche Intelligenz im Gesundheitswesen aufgeführt, die im Zeitraum von 2018 bis 2020 veröffentlicht wurde. Die drei als relevant erachteten Literaturanalysen decken zusammen einen Publikationszeitraum von 2010 bis 2020 ab und können zum gegenwärtigen Zeitpunkt – 2 Jahre später – als aktuell angesehen werden. Daher dienen die vorliegenden systematischen Literaturanalysen als guter Ausgangspunkt für eine Untersuchung der dort verwendeten KI-Kategorienansätze und Ausprägungen.

Tabelle 3 Überblick über systematische Literaturanalysen zum Thema KI im Gesundheitswesen

ID	Literaturverweise	Titel	Typ	Input	Versorgungs- säule	Fachgebiet	Nutzer	Technologie	Output
L03	Atanasov, Gauthier und Lopes (2018)	Applications Of Artificial Intelligence Technologies In Healthcare: A Systematic Literature Review	Systematic Literature Review			✓			
L07	Ben-Israel, Jacobs, Casha, Lang, Ryu, de Lotbiniere-Bassett und Cadotte (2020)	The impact of machine learning on patient care: A systematic review	Systematic Literature Review	✓		✓			
L60	Wolff, Pauling, Keck und Baumbach (2020)	The Economic Impact of Artificial Intelligence in Health Care: Systematic Review	Systematic Literature Review				✓		

Vor diesem Hintergrund wurde ausgehend von den im vorhergehenden Abschnitt definierten Forschungsfragen Suchbegriffe abgeleitet, die auf deren Beantwortung abzielen. Zu diesem Zweck wurden geeignete englische Begriffe und Phrasen ausgewählt und über Boolesche Operatoren zu geeigneten Schlüsselwortsuchen kombiniert, welche wiederum an die unterschiedliche Syntax der unterschiedlichen Literaturdatenbanken angepasst wurden und durch Testläufe in diesen Literaturdatenbanken schrittweise verfeinert wurden. Um die teilweise großen Suchergebnismengen der Testsuchläufe auf eine handhabbare Größe zu reduzieren, wurde die Suche jeweils auf das Datenbankfeld Titel beschränkt, so dass ausschließlich Suchtreffer angezeigt wurden, bei denen eine geeignete Kombination der Suchbegriffe im Titel gefunden wurde. Der in der Schlüsselwortsuche verwendete Suchausdruck für bestehende Kategorisierungen, Systematisierungen sowie Taxonomien (RQ 1.1) lässt sich vor diesem Hintergrund in einer generischen Syntax wie folgt darstellen:

*[title: (((**"AI"** OR **"Artificial intelligence"**) AND (**"categori*"** OR **"systemat*"**) AND (**"health*"** OR **"care"** OR **"medic*"**)))]*

Der final in der Schlüsselwortsuche verwendete Suchausdruck zur Ableitung der Kategorien und Ausprägungen (RQ 1.2 + 1.3) lautet wie folgt in einer generischen Syntax:

*title: ((**"artificial intelligence"** OR **"AI"** OR **"machine learning"** OR **"ML"** OR **"deep learning"** OR **"DL"** OR **"neural network"** OR **"NN"**) AND (**"medic*"** OR **"clinic*"** OR **"health*"** OR **"care"**))*

Insgesamt wurden drei Online-Literaturdatenbanken durchsucht, welche aus den Bereichen Medizin, Psychologie und Technik ausgewählt wurden (vgl. Tabelle 4).

Tabelle 4 Übersicht über die verwendeten Literaturdatenbanken

Literaturdatenbank	MED	PSYC	TECH
Web of Science	✓	✓	✓
Scopus	✓	✓	✓
PubMed	✓	✓	
<i>Legende: MED = Medizin; PSYC = Psychologie; TECH = Technik</i>			

Der Suchzeitpunkt und die Suchfilter, die bei der ersten Stichwortsuche in den o.g. drei Online-Literaturdatenbanken verwendet wurden, sind in Tabelle 5 zusammengefasst:

Tabelle 5 Überblick über Suchzeitpunkt und Suchfilter

Merkmal	Beschreibung
Suchzeitpunkt	Die initiale Schlüsselwortsuche wurde im September/Oktober 2021 in den drei genannten Online-Literaturdatenbanken mit den zu diesem Zeitpunkt jeweils verfügbaren Suchalgorithmen und Datenbankinhalten durchgeführt. Die Durchführung derselben Suche kann zu einem späteren Zeitpunkt aufgrund veränderter Algorithmen und Datenbankinhalte zu unterschiedlichen Suchergebnissen führen.
Publikationszeitraum	Je Online-Literaturdatenbank wurde der komplette verfügbare Datenbestand ab dem Jahr 2010 durchsucht.

Durchsuchte Datenbankfelder	Die Suche wurde auf das Titelfeld beschränkt, da irrelevante Publikationen zu Beginn der Relevanzeinschätzung ohnehin auf Basis von Titel und Kurzfassung herausgefiltert wurden.
Sprache	Es wurde ausschließlich Literatur in englischer und deutscher Sprache berücksichtigt, wobei jedoch mit englischen Suchbegriffen gearbeitet wurde.
Publikationstypen	Von der Suche wurden – sofern von der Online-Literaturdatenbank unterstützt – die Publikationstypen Clinical Trials, Randomized Controlled ausgeschlossen, da sich diese Ergebnisse zu spezifisch für die zu erarbeitende Übersicht darstellten.

In Ergänzung zu den Online-Literaturdatenbanken wurde das Internet mit Hilfe der Suchmaschinen Google und Google Scholar als sonstige Literaturquellen nach weiteren nicht-/wissenschaftlichen Publikationen und fehlenden Volltexten durchsucht.¹⁶

Abbildung 1 stellt den durchgeführten Prozess der Literatursuche und Literaturauswahl dar.

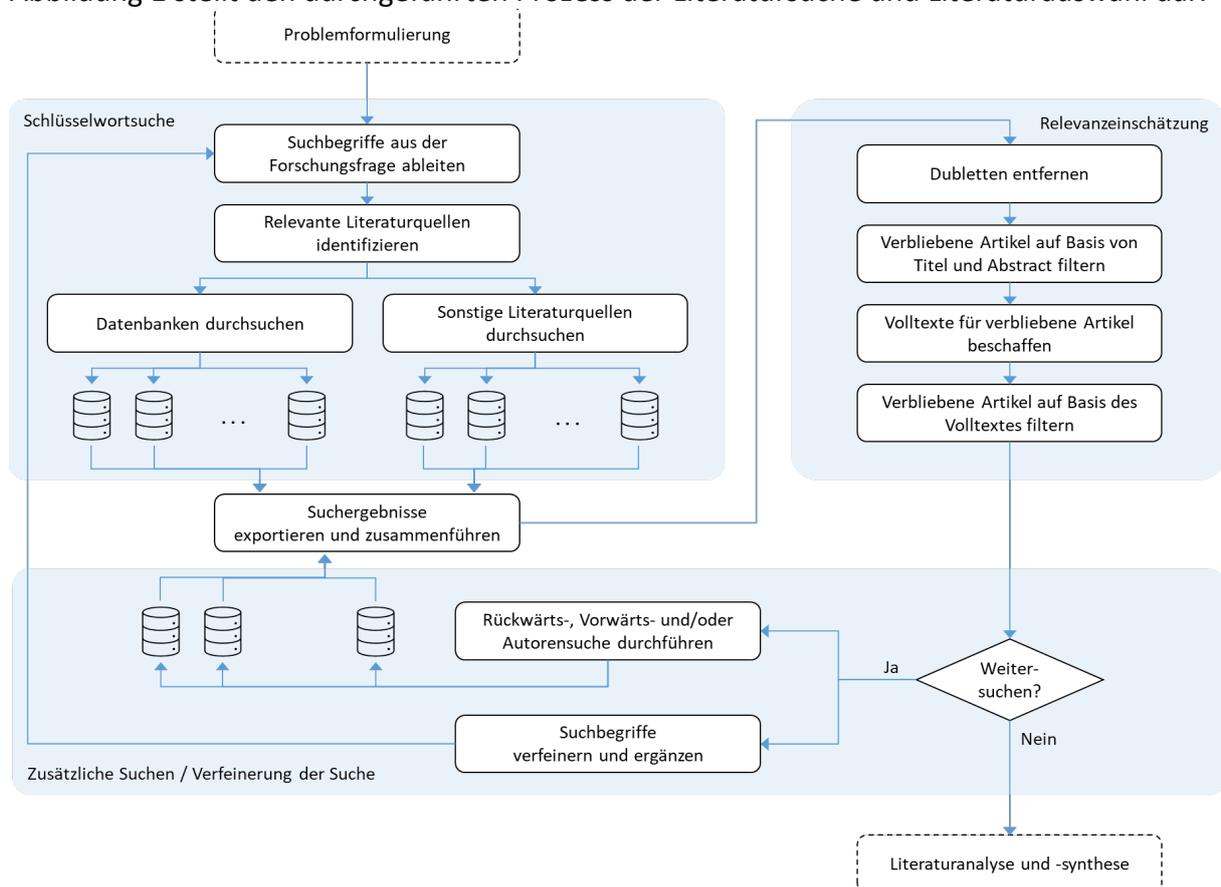


Abbildung 1 Literatursuche und -auswahlprozess (eigene Darstellung in Anlehnung an Schryen (2015), Wolfswinkel, Furtmueller und Wilderom (2013), Zhang Babar und Tell (2011) und vom Brocke et. al. (2009)

Durch die Bündelung verschiedener Suchstrategien und Literaturquellen wurden insgesamt 8.254 Publikationen ermittelt, die sukzessive auf 61 relevante Publikationen reduziert wurden.

¹⁶ Rothstein und Hopewell (2009), Kitchenham und Charters (2007) und Petticrew und Roberts (2006)

Tabelle 6 verdeutlicht, welchen wesentlichen Beitrag die Einbeziehung zusätzlicher Suchstrategien und Literaturquellen, die über eine reine Stichwortsuche in Online-Literaturdatenbanken hinausgehen, zur Ganzheitlichkeit der Literaturanalyse leistet.

Tabelle 6 Zusätzliche Suchstrategien und Literaturquellen

Suchstrategien und Literaturquellen	Publikationen insgesamt	Relevante Publikationen
Schlüsselwortsuche	8.254	41
Rückwärtssuche	1	1
Vorwärtssuche	0	0
Autorensuche	3	3
Sonstige Literaturquellen	16	16
Summe	8.274	61

Abbildung 2 veranschaulicht, wie groß der Dublettenanteil nach Zusammenführung der ersten Stichwortsuchen in den verschiedenen Online-Literaturdatenbanken war und wie sich für jede Phase der Relevanzabschätzung relevante und irrelevante Publikationen auf die verschiedenen Suchstrategien und Literaturquellen verteilen:

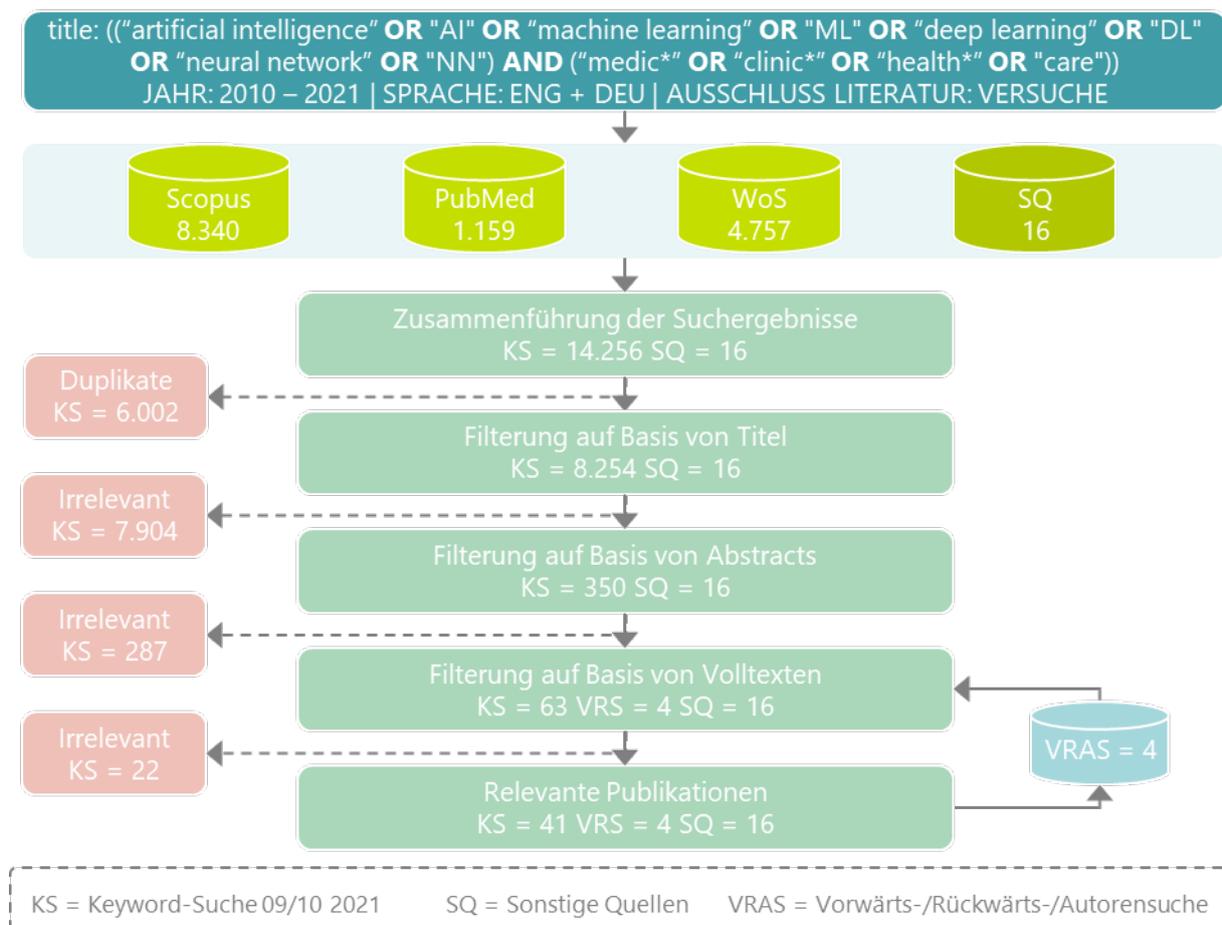


Abbildung 2 Überblick über Dubletten und Relevanzabschätzung

Die während der Relevanzeinschätzung verwendeten Ein- und Ausschlusskriterien sind in Tabelle 7 detaillierter dargestellt:

Tabelle 7 Ein- und Ausschlusskriterien für die Literatursuche

Einschlusskriterium	Beschreibung
Kategorisierungsansätze von KI in der Versorgung	Die Publikation fokussiert mögliche Kategorisierungsansätze von KI in der Versorgung und beschreibt Ausprägungen mind. einer Kategorie (<i>Input, Versorgungssäule, Fachgebiete, Nutzer, Technologie, Output</i>)
Ausschlusskriterium	Beschreibung
Ungültige Sprache	Die Publikation liegt in einer anderen Sprache als Englisch oder Deutsch vor.
Ungültige Studienart	Die Publikation stellt die Ergebnisse von klinischen Versuchen dar (z.B. trial, assessment, performance) und verfehlt damit das Ziel überblicksartiger Literaturquellen.
Nicht zielführender Fokus	Die Publikation bildet z.B. keine interdisziplinären Kategorien und Ausprägungen der KI in der Versorgung ab und stellt nur einseitige Bereiche und Akteure dar. Weiterhin werden komplexe KI-Techniken (Algorithmen, Entscheidungsbäume, Bayessche Netze etc.), spezifische Krankheitsphänomene (z.B. COVID) und Fachgebiete (Veterinärmedizin, Psychologie, Notarztwesen etc.) zu stark fokussiert, was im Sinne einer überblicksartigen Zielstellung zum Ausschluss der Publikationen führt.

3.2.3 Analyse und Interpretation der relevanten Literatur

Im nächsten Schritt wurden die 61 als relevant erachteten Publikationen gesammelt, um sie zu analysieren und zu interpretieren. Dabei wurde eine konzeptzentrierte Ergebnisdarstellung verwendet, in der die Ergebnisse der relevanten Literatur in verschiedenen Konzeptmatrizen erfasst, strukturiert und kategorisiert wurden und so verschiedene Arbeiten zueinander in Beziehung gesetzt werden konnten.¹⁷

Zunächst wurden die relevant erachteten Artikel auf einer ersten Stufe kategorisiert und verglichen. Vor diesem Hintergrund sind in zum einen je Publikation der zugehörige Publikationstyp aufgeführt. Zum anderen ist dort ersichtlich, welche Aspekte der Kategorisierung von KI in der existierenden Literatur untersucht wurden, indem je Publikation angegeben ist, welche Kategorie näher beschrieben wird: Input (6), Versorgungssäule (14), Fachgebiete (7), Nutzer (7), Technologieklasse (7), Output (4). Hierbei ist zu berücksichtigen, dass Mehrfachnennungen möglich sind, da dieselbe Publikation gleichzeitig mehrere Kategorien abdecken kann. Als

¹⁷ Vgl. Schryen (2015), Wolfswinkel, Furtmueller und Wilderom (2013), vom Brocke et al. (2009), Levy und Ellis (2006) und Webster und Watson (2002)

Ergebnis wurde aus diesem Schritt der in Abbildung 3 dargestellte 1. Entwurf eines Kategorisierungsinstrumentes für KI in der Versorgung abgeleitet.

		Kategorien					
		Input	Versorgung	Fachgebiet	Nutzer	Technologie	Output
Ausprägungen	Bildgebungsdaten	Prävention	Chirurgie	Patienten	Bildererkennung und -verstehen	Umfassendere Betreuung	
	Audiodaten	Diagnostik	Dermatologie	Ärzte	Datenmanagement und -analyse	Verbesserte Beratung	
	Zeitreihendaten	Therapie	Innere Medizin	Angehörige von Patienten	Mensch-Maschine-Interaktion und Assistenzsysteme	Präzisere Diagnose	
	Daten aus Untersuchung, Labor, Demografie	Rehabilitation	Intensivmedizin	Medizinisches/-technisches Fachpersonal	Robotik und autonome Systeme	Personalisierte Behandlung	
	Sonst. klinische Daten (handschr. Notizen, Begegnungen)	Pflege	Neurologie	Pflegerisches Fachpersonal	Sensorik und Kommunikation	Verbesserte (Behandlungs-) Abläufe	
	"-omics" Daten		Öffentliche Gesundheit	Forscher	Sprach- und Textverstehen	Verbesserte (Versorgungs-) Forschung & Entwicklung	
			Pathologie	Krankenkassen	Virtuelle und augmented Realität		
			Pflege				
			Pharmakologie				
			Psychiatrie				
		Radiologie					
		Urologie					

Abbildung 3 1. Entwurf eines Kategorisierungsinstrumentes für KI in der Versorgung

4. Experteninterviews zur Validierung der Kategorisierung von KI in der Versorgung

4.1 Methodik

Um das aus dem Literature Review abgeleitete Kategorisierungssystem im Rahmen von Experteninterviews mit Experten aus der Gesundheitswirtschaft, Medizin, KI-Forschung, Medizintechnik und Versorgungsforschung zu evaluieren und um exemplarische Vertreter von KI-Systemen zu ergänzen, wurde zunächst ein semistrukturiertes Interview konzipiert, mit einem KI-Experten getestet und im Anschluss leicht modifiziert. Das modifizierte Interview, das für die tatsächliche Befragung zum Einsatz kam, deckte folgende Bereiche ab:

- Demografika (Geschlecht, Alter, Beruf, spezifisches KI-Forschungsfeld, Beschäftigungsdauer mit KI, Einschätzung des eigenen KI-Expertenwissens auf einer 5er-Likert-Skala)
- Abgleich der Definition von KI des Probanden mit der im Projekt verwendeten Definition
- Freie Sammlung von KI-Versorgungsbeispielen inkl. Status (Entwicklung vs. Zulassung vorliegend), Vor- und Nachteilen sowie Chancen und Risiken des jeweiligen Beispiels.
- Vorstellung des auf Basis der Literaturanalyse entwickelten Kategorisierungsinstrumentes und erste Eindrücke seitens der Probanden
- Kategorisierung eines selbst gewählten KI-Versorgungsbeispiels anhand des Kategorisierungsinstrumentes inkl. Evaluation
- Versuch der Erklärung der eingangs genannten KI-Versorgungsbeispiele für einen KI-Neuling (z.B. Oma)
- Einschätzung der eigenen Nutzungswahrscheinlichkeit der KI-Versorgungsbeispiele auf einer 5er-Likert-Skala
- Evaluation des Kategorisierungssystems hinsichtlich der Aspekte Vollständigkeit, Verständlichkeit und Interdisziplinarität auf einer 5er-Likert-Skala

Neben dem Interviewablauf wurde auch das aus dem Literature Review abgeleitete Kategorisierungssystem mit einem KI-Experten sowie den KollegInnen des IPG hinsichtlich Vollständigkeit und Verständlichkeit getestet und entsprechend modifiziert bzw. als zusätzliche abzufragende Bereiche in das semistrukturierte Interview aufgenommen. Abbildung 4 zeigt das modifizierte Kategorisierungssystem, das die Grundlage für die durchgeführten Experteninterviews bildete:

		Kategorien				
Ausprägungen	Input	Versorgung	Fachgebiet	Nutzer	KI-basierte Technologie	Output
	Bilddaten	Prävention	Chirurgie	Patienten	Bildererkennung und -verstehen	Umfassendere Betreuung
	Audiodaten	Diagnose	Dermatologie	Ärzte	Big Data-Analyse	Verbesserte Beratung
	Zeitreihendaten	Therapie	Innere Medizin	Angehörige von Patienten	Mensch-Maschine-Interaktion und Assistenzsysteme	Präzisere (Ausschluss-) Diagnose
	Daten aus Untersuchung, Labor, Demografie	Rehabilitation	Intensivmedizin	Medizinisches/-technisches Fachpersonal	Robotik und autonome Systeme	Personalisierte Behandlung
	Sonst. klinische Daten (handschr. Notizen, Begegnungen)	Pflege	Neurologie	Pflegerisches Fachpersonal	Biosignal- / Vitalwert- / Sensorikanalyse	Verbesserte (Behandlungs-) Abläufe
	"-omics" Daten		Öffentliche Gesundheit	Forscherin	Sprach- und Textverstehen	Verbesserte (Versorgungs-) Forschung & Entwicklung
	Vitaldaten		Pathologie	Krankenkasse	Virtuelle und augmented Realität	Verbesserte Detektion von Anomalien, Auffälligkeiten
	Anamnesedaten		Pflege	Krankenhaus		Verbesserte Dokumentation
			Pharmakologie	Sonstige stationäre Einrichtungen		
		Psychiatrie				
		Radiologie				
		Urologie				
		Onkologie				

Abbildung 4 Kategorisierungssystem als Grundlage für Experteninterviews

4.2 Ergebnis

4.2.1 Demografika

Im Zeitraum vom 27.10.-07.12.21 wurden insgesamt sieben Experteninterviews durchgeführt, wobei ein Interview in Präsenz stattfand. Alle weiteren Interviews wurden aufgrund der neuerlichen Corona-Einschränkungen als Online-Interviews über Microsoft Teams durchgeführt. Für die Auswertung der auf Tonband bzw. Video aufgezeichneten Interviews wurde eine qualitative Inhaltsanalyse in Anlehnung an Mayring (2015) durchgeführt und insgesamt 772 Einzelcodes identifiziert.

Tabelle 8 zeigt eine demografische Übersicht über die TeilnehmerInnen der KI-Experteninterviews.

Tabelle 8 Soziodemografische Angaben

	Mittelwert / Häufigkeit	Minimum/Maximum	N
Alter	45,53 (SD=10,90)	31/62	7
Geschlecht: weiblich	3 (42,85%)		7
Geschlecht: männlich	4 (57,14%)		7
Beschäftigung mit KI in Jahren	10,71 (SD=10,62)	3/30	7
Selbsteinschätzung KI-Expertenwissen (1= kein Wissen; 5= Experte)	2,86 (SD=,69)	2/4	7

Anmerkung: Angegeben sind bei numerischen Werten Mittelwerte inkl. des kleinsten bzw. größten Werts, bei anderen Werten Häufigkeiten. SD entspricht der Standardabweichung des Mittelwertes. N gibt die Anzahl der Personen an, die die Frage beantwortet haben. Prozentuale Häufigkeiten sind jeweils bezogen auf die Gesamtstichprobe von 7 Teilnehmern berechnet.

In Bezug auf den beruflichen Hintergrund war die Teilnehmergruppe breit aufgestellt und beinhaltete neben Informatikern auch einen Innovationsmanager im Gesundheitsbereich sowie eine Physiotherapeutin mit Schwerpunkt Pflege sowie einen Facharzt für Anästhesiologie und Intensivmedizin.

Auffällig in Bezug auf die Selbsteinschätzung des KI-Expertenwissens war, dass – obwohl sich die TeilnehmerInnen im Mittel bereits 10,71 Jahre mit KI beschäftigten – sie ihr Expertenwissen eher im Mittelfeld einstuften (1= kein Wissen; 5 = Experte). Dies wurde häufig damit begründet, dass man trotz eines großen Wissens und praktischer Erfahrung mit KI in der Versorgung selbst keine KI programmiere bzw. die Programmierung zeitlich bereits länger zurückläge.

4.2.2 Definitionen von KI

Obwohl die spezifischen Definitionen von KI – der Literatur folgend – sehr unterschiedlich ausfielen, lag ein Definitionsfokus auf selbstlernenden Systemen und war damit enger gefasst als die im Projekt KI-BA verwendete, weiter gefasste Definition von KI als jeder Art von regel-, daten- oder optimierungsbasiertem System. Die in den Interviews genannten Definitionen reichten von KI als technischer Unterstützung zur Entscheidungsfindung, einem Nachbau menschlicher Intelligenz mit Maschinen, Methoden des maschinellen Lernens, selbstlernenden Algorithmen bis hin zu einer Unmöglichkeit der Definition von KI („KI ist alles und nichts“).

4.2.3 Erste Eindrücke zum Kategorisierungssystem

Der erste Eindruck des Kategorisierungssystems wurde von den TeilnehmerInnen als sinnhaft, umfassend, interessant, relativ vollständig und inspirierend in Bezug auf weitere Anwendungen geschildert. In Bezug auf die Darstellungsart und die einzelnen Kategorien und ihre Ausprägungen wurden jedoch auch zahlreiche Ergänzungen und Verbesserungsvorschläge genannt, die nachfolgend aufgelistet sind:

- **Darstellung:** Bezüglich der Darstellung wurde mehrfach angemerkt, dass die Ausprägungen nicht überschneidungsfrei seien. So können in ein System durchaus mehrere Input-Ausprägungen einfließen, mehrere Nutzer, Versorgungs- und Fachgebiete betroffen sein und auch mehrere KI-basierte Technologien zur Anwendung kommen. Vor diesem Hintergrund wurde angeregt, die gewählte Darstellungsart des morphologischen Kastens ggf. durch eine Darstellungsart zu ersetzen, die mehrere Kombinationsmöglichkeiten zulässt. Vorgeschlagen wurde hier eine sternenförmige statt einer linearen Anordnung. Vor dem Hintergrund des Kategorisierungssystems als Auswahlhilfe für KI-Beispiele für die spätere Akzeptanzstudie wurde letztlich auf eine Änderung der Darstellung verzichtet, da zu viele Überschneidungen zu nur bedingt trennscharfen KI-Beispielen führen würden.
- **Kategorie „Input“:** Hier wurde angeregt, die Kategorie „Input“ um die Ausprägungen Anamnesedaten (Ergebnis aus Pretest), Fachliteratur/Publicationen und GKV-Routinedaten (Abrechnungsdaten) zu ergänzen. Die Ergänzungen wurden für die finale Version des Kategorisierungssystems umgesetzt. Nicht umgesetzt wurden die Anregungen, Daten aus der Patientenakte bzw. Daten aus DiGAs zu ergänzen, da diese Daten in ihrer ursprünglichen Form bereits im Kategorisierungssystem vertreten sind und die Patientenakte bzw. die DiGAs primär ein neues Speichermedium darstellen.
- **Kategorie „Versorgung“:** Hier wurde angeregt, die Kategorie „Versorgung“ um die Ausprägung Monitoring als Vor-/Nachsorge-Kreislauf zu ergänzen. Dieser Vorschlag wurde jedoch nicht übernommen, da Prävention (= Vorsorge) bzw. Rehabilitation und Pflege (= Nachsorge) bereits entsprechende Teilkategorien des Monitoring-Prozesses im Kategorisierungssystem vertreten sind.
- **Kategorie „Fachgebiet“:** Hier wurde angeregt, die Kategorie „Fachgebiet“ um die stark interdisziplinär ausgerichteten Ausprägungen Onkologie (Ergebnis aus Pretest), Notarztwesen und Sportmedizin zu ergänzen und Psychologie als zusätzliche Ausprägung neben Psychiatrie einzufügen. Die Ergänzungen wurden für die finale Version des Kategorisierungssystems umgesetzt. Auf die Umsetzung der Anregung, dass die

Ausprägung „Innere Medizin“ zu grob sei und es hier die Teilgebiete aufgelistet werden sollten, wurde verzichtet, da dies der Übersichtlichkeit des Kategorisierungssystems entgegenstehen würde.

- **Kategorie „Nutzer“:** Hier wurde angeregt, die Kategorie „Nutzer“ um die Ausprägungen Apotheke, Heilmittelerbringen (Physio-, Ergo-, Logopädie), Hilfsmittelverordner, Krankenhausverwaltungen, Rettungsdienst und Versicherungen zu ergänzen. Die Ergänzungen wurden für die finale Version des Kategorisierungssystems umgesetzt. Da ferner angemerkt wurde, dass Berufsgruppen und Einrichtungen vermischt seien, wurden die Formulierungen für die finale Version so angepasst, dass jeweils Personen/Berufsgruppen adressiert werden.
- **Kategorie „KI-basierte Technologie“:** Hier wurde angeregt, die Kategorie „KI-basierte Technologie) um Explainable AI zu ergänzen. Ferner wurde angeregt, die Ausprägung Sensorik in Big Data-Analyse und Biosignalanalyse zu trennen (Ergebnis aus Pretest). Die Ergänzungen wurden für die finale Version des Kategorisierungssystems umgesetzt. Ebenfalls angemerkt wurde, dass die „KI-basierte Technologie“ eher das umgebende System statt die eigentliche Technologie (z.B. Bayessche Netze) beschreibt. Auf eine Umbenennung bzw. weitere Differenzierung bzgl. der eigentlichen Technologie wurde verzichtet, da diese oftmals nicht bekannt bzw. von den Entwicklern offengelegt wird.
- **Kategorie „Output“:** Hier wurde angeregt, die Kategorie „Output“ um frühzeitige Intervention sowie verbesserte Behandlung zu ergänzen. Die Ergänzungen wurden für die finale Version des Kategorisierungssystems umgesetzt. Nicht umgesetzt wurde der Vorschlag, die Kategorie umzubenennen in „KI-basierter Vorschlag für Output“, da dies den Output zu stark auf eine bestimmte Art von KI einschränkt, andere Arten dagegen völlig ausblendet (z.B. KI-Variante, die selbstständig entscheiden und handeln).

4.2.4 Kategorisierung eigener Beispiele im Kategorisierungssystem

Insgesamt wurden von den TeilnehmerInnen 19 verschiedene KI-Beispiele und anhand des Kategorisierungssystems kategorisiert. Abbildung 5 zeigt eine Übersicht der KI-Beispiele sowie ihrer Klassifizierung.

Titel	Input	werden mittels	KI-basierte Technologie	durch	Nutzer	der	Fachgebiet	zur	Versorgung	aufbereitet für	Output
Multimodale Zeitreihendaten (Fitnessshirt-Monitoring)	Zeitreihendaten	werden mittels	Biosignal-/Vitalwert-/Sensorikanalyse	durch	Ärzte	der	Inneren Medizin	zur	Diagnostik	aufbereitet für	präzisere (Ausschluss-)Diagnosen.
KI-basierte Schmerzanalyse in der Pflege	Bilddaten, Zeitreihendaten sowie sonstige klinische Daten	werden mittels	Bildererkennung und -verstehen und Biosignal-/Vitalwert-/Sensorikanalyse	durch	pflegerisches und medizinisches/technisches Fachpersonal	der	Intensivmedizin und Pflege	zur	Diagnostik	aufbereitet für	verbesserte Detektionen von Anomalien und Auffälligkeiten, verbesserte Dokumentationen, präzisere (Ausschluss-) Diagnosen.
KI-basierte Detektion von Läsionen / Polypen	Bilddaten	werden mittels	Bildererkennung und -verstehen	durch	Ärzte	der	Inneren Medizin	zur	Prävention	aufbereitet für	verbesserte Detektionen von Anomalien und Auffälligkeiten.
AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	Bilddaten, Audiodaten, Zeitreihendaten sowie sonstige klinische Daten	werden mittels	Bildererkennung und -verstehen, Biosignal-/Vitalwert-/Sensorikanalyse, Sprach- und Textverstehen sowie Virtual und Augmented Reality	durch	pflegerisches Fachpersonal	der	Pflege, Dermatologie sowie Inneren Medizin	zur	Prävention, Therapie und Pflege	aufbereitet für	umfassendere Betreuungen, verbesserte, personalisierte Behandlungen, verbesserte (Behandlungs-) Abläufe, verbesserte Detektionen von Anomalien und Auffälligkeiten sowie verbesserte Dokumentationen.
Antibiotikagabe im Krankenhaus	Daten aus Untersuchungen, Laboren, Demografien	werden mittels	Biosignal-/Vitalwert-/Sensorikanalyse	durch	Ärzte	der	Intensivmedizin	zur	Therapie	aufbereitet für	verbesserte (Behandlungs-) Abläufe
Genetischspezifizierte Antibiotikagabe im Krankenhaus	Anamnesedaten, Daten aus Untersuchungen, Laboren, Demografien sowie Omics-Daten	werden mittels	Biosignal-/Vitalwert-/Sensorikanalyse	durch	Ärzte	der	Intensivmedizin	zur	Therapie	aufbereitet für	verbesserte (Behandlungs-) Abläufe
Ultraschallbasierte Flüssigkeitsgabe im Notarzttwesen	Bilddaten	werden mittels	Bildererkennung und -verstehen	durch	Ärzte	des	Notarzttwesens	zur	Diagnostik und Therapie	aufbereitet für	präzisere (Ausschluss-) Diagnosen.
Frühzeitige Erkennung von Unterzuckerung auf der Intensivstation	Daten aus Untersuchungen, Laboren, Demografien	werden mittels	Mensch-Maschine-Interaktionen & Assistenzsysteme sowie Biosignal-/Vitalwert-/Sensorikanalyse	durch	Ärzte und pflegerisches Fachpersonal	der	Intensivmedizin	zur	Therapie	aufbereitet für	präzisere (Ausschluss-) Diagnosen und verbesserte (Behandlungs-) Abläufe.
Frühzeitige Erkennung von Unterzuckerung für den Privatgebrauch	Daten aus Untersuchungen, Laboren, Demografien	werden mittels	Mensch-Maschine-Interaktionen & Assistenzsysteme sowie Biosignal-/Vitalwert-/Sensorikanalyse	durch	Patienten und Angehörige der Patienten	(der	Intensivmedizin)	zur	Therapie	aufbereitet für	präzisere (Ausschluss-) Diagnosen und verbesserte (Behandlungs-) Abläufe.
KI-basierte Erkennung von Brustkrebs	Bilddaten	werden mittels	Bildererkennung und -verstehen	durch	Ärzte und medizinisches/technisches Fachpersonal	der	Radiologie	zur	Prävention	aufbereitet für	präzisere (Ausschluss-) Diagnosen.
GKV-Routinedaten für Parkinson-Sturz-Prävention	Routinedaten der GKV	werden mittels	Big Data-Analyse	durch	Personal der Krankenkassen	(der	Neurologie)	zur	Prävention	aufbereitet für	verbesserte Beratungen.
Longitudinale Datenauswertung / Krankheitsprädiktion	Bilddaten, Zeitreihendaten und Daten aus Untersuchungen, Laboren, Demografien	werden mittels	Bildererkennung und -verstehen, Biosignal-/Vitalwert-/Sensorikanalyse sowie Big Data-Analyse	durch	Ärzte	der	Inneren Medizin	zur	Therapie	aufbereitet für	verbesserte, personalisierte Behandlungen und verbesserte (Behandlungs-) Abläufe.
Verletzungsdiagnose paralympischer Athleten	Sonstige klinische Daten	werden mittels	Sprach- und Textverstehen	durch	medizinisches/technisches Fachpersonal	der	Sportmedizin	zur	Prävention	aufbereitet für	umfassendere Betreuungen und verbesserte Detektionen von Anomalien & Auffälligkeiten.
Autismustherapie	Bilddaten und Zeitreihendaten	werden mittels	Bildererkennung und -verstehen, Mensch-Maschine-Interaktion & Assistenzsysteme, Biosignal-/Vitalwert-/Sensorikanalyse sowie Big Data-Analyse	durch	medizinisches/technisches Fachpersonal und Angehörige von Patienten	der	Psychologie	zur	Therapie	aufbereitet für	verbesserte (Behandlungs-) Abläufe.
Blutbildanalyse Hiamacam	Bilddaten	werden mittels	Bildererkennung und -verstehen	durch	Ärzte und medizinisches/technisches Fachpersonal	der	Inneren Medizin	zur	Diagnostik	aufbereitet für	präzisere (Ausschluss-) Diagnosen.
App für Kardiopatienten	Diskrete Vitaldaten, kontinuierliche Zeitreihendaten sowie Anamnesedaten	werden mittels	Biosignal-/Vitalwert-/Sensorikanalyse	durch	Patienten und Ärzte	der	Inneren Medizin	zur	Diagnostik und Therapie	aufbereitet für	präzisere (Ausschluss-) Diagnosen, verbesserte, personalisierte Behandlungen, verbesserte (Behandlungs-) Abläufe und verbesserte Detektionen von Anomalien & Auffälligkeiten.
Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	Zeitreihendaten	werden mittels	Biosignal-/Vitalwert-/Sensorikanalyse	durch	Ärzte	der	Neurologie	zur	Diagnostik und Therapie	aufbereitet für	verbesserte, personalisierte Behandlungen sowie verbesserte Detektionen von Anomalien & Auffälligkeiten.
Digitale Pathologie (medizinische Bildanalyse)	Bilddaten	werden mittels	Bildererkennung und -verstehen	durch	Ärzte	der	Pathologie	zur	Diagnostik	aufbereitet für	präzisere (Ausschluss-) Diagnosen.
Leitlinienmodellierung / longitudinale Modellierung	Zeitreihendaten	werden mittels	Big Data-Analyse	durch	Ärzte	der	Onkologie	zur	Diagnostik und Therapie	aufbereitet für	präzisere (Ausschluss-) Diagnosen und verbesserte, personalisierte Behandlungen.

Abbildung 5. Übersicht aller KI-Beispiele und Einordnung ins Kategorisierungssystem.

Nachfolgend werden die einzelnen genannten KI-Beispiele mit ihrem Entwicklungsstand (Entwicklung vs. Zulassung) sowie der von den TeilnehmerInnen angeführten Chancen und Risiken und der eigenen Nutzungswahrscheinlichkeit dargestellt. Inhaltlich ähnliche Beispiele (z.B. Weiterentwicklung eines Beispiels für einen anderen Anwendungskontext) sind dabei zusammengefasst:

- **Multimodale Zeitreihendaten (Fitnessshirt-Monitoring):**
 - **Entwicklungsstand:** In der Entwicklung
- **KI-basierte Schmerzanalyse in der Pflege:**
 - **Entwicklungsstand:** In der Entwicklung
 - **Chancen und Risiken:** Aktuell Mangel schmerzbasierter Bilder; Möglichkeit des Home-Self-Assessments als App für Patienten; Entlastung der Pflege; Objektivierung der Dokumentation
- **KI-basierte Detektion von Läsionen/Polypen:**
 - **Entwicklungsstand:** FDA-Zulassung, aber noch kein täglicher Einsatz
 - **Chancen und Risiken:** Gefahr des Übersehens von Dingen aufgrund zu starken Verlasses auf die KI; Entlastung von (jungen) Ärzten
- **AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses:**
 - **Entwicklungsstand:** Wünschenswerte Entwicklung
 - **Chancen und Risiken:** Bessere Versorgung von Patienten/Bewohnern; für Personal Sicherheit über optimale Betreuung; Einsatz weniger geschulten Personals mit KI-Unterstützung möglich (ggf. Anpassung gesetzlicher Rahmenbedingungen nötig); Optimierung von Personal und Versorgung durch zweiten Blick; Risiko der Personaleinsparung durch KI-Einsatz; Frage der Kosten;

- **Eigene Nutzungswahrscheinlichkeit:** Auf einer Skala von 1 (= Nutzung sehr unwahrscheinlich) bis 5 (= Nutzung sehr wahrscheinlich) sowohl aus Arzt- als auch als Patientensicht: „Wenn sie das kann, was ich von ihr erwarte bei 5, wenn sie es nicht kann, bei 1.“
- **(Genetisch spezifizierte) Antibiotikagabe im Krankenhaus:**
 - **Entwicklungsstand:** In der Entwicklung bzw. Wunsch im Falle von genetisch spezifizierter Antibiotikagabe
 - **Chancen und Risiken:** Risiko eines Behandlungsfehlers wegen fehlerhafter Dateneingabe des Personals; kein eigenständiges Denken des Personals mehr aufgrund von Verlass auf KI; weniger Gabe von Breitbandantibiotika nötig, dadurch weniger Resistenzen beim Patienten; weniger Medikamentinteraktionen nötig (Apothekervisite); verbesserte Anhaltspunkte für Apotheken zur Beschaffung von Therapeutika; zielsicherer Einsatz von Antibiotika und damit mehr Sicherheit für Patienten; Unterstützung des Personals bei der Behandlung; Notwendigkeit der Anlage und Pflege von Datenbanken; Risiko der Datensicherheit; ggf. fehlende Akzeptanz
 - **Eigene Nutzungswahrscheinlichkeit:** Auf einer Skala von 1 (= Nutzung sehr unwahrscheinlich) bis 5 (= Nutzung sehr wahrscheinlich) sowohl aus Arzt- als auch aus Patientensicht Nutzung jeweils sehr wahrscheinlich (5).
- **Ultraschallbasierte Flüssigkeitsgabe im Notarztwesen:**
 - **Entwicklungsstand:** Wünschenswerte Entwicklung
 - **Chancen und Risiken:** Risiko des zu starken Verlasses auf die KI; Notwendigkeit von Schulungen (auch seitens der Industrie) zur Vermeidung von eventuellen Anwendungsfehlern seitens des Personals; Risiko einer Fehltherapie/-diagnose bei fehlerhafter Anwendung; Risiko eventueller Berührungsgänge seitens des älteren Personals; Verminderung des akuten Zeitstresses im Krankenhaus; optimierte (ggf. lebensrettende) Diagnose- und Therapiechancen für Patienten; Sicherheit bzgl. richtiger Behandlung, insb. für unqualifiziertes Personal;
 - **Eigene Nutzungswahrscheinlichkeit:** Auf einer Skala von 1 (= Nutzung sehr unwahrscheinlich) bis 5 (= Nutzung sehr wahrscheinlich) aus Patientensicht Nutzung wahrscheinlich (4) bzw. aus Arztsicht Nutzung sehr wahrscheinlich (5), jeweils unter der Bedingung der Funktionalität im Sinne von geringer Fehlerquote.
- **Frühzeitige Erkennung von Unterzuckerung auf der Intensivstation bzw. für den Privatgebrauch:**
 - **Entwicklungsstand:** Im Krankenhaus in Anwendung; für den Privatgebrauch wünschenswerte Entwicklung
 - **Chancen und Risiken:** Unterstützung bei personellem Notstand; schnelle Reaktion des Personals auf den Patienten
- **KI-basierte Erkennung von Brustkrebs:**
 - **Entwicklungsstand:** Im Einsatz als zugelassenes Medizinprodukt
 - **Chancen und Risiken:** Risiko des völligen Verlassens auf die Entscheidung der KI; Risiko der KI als Black Box mit Notwendigkeit der verständlichen Kommunikation sowohl gegenüber Leistungserbringern als auch Patienten; Risiko eines fehlenden Vertrauens/fehlender Akzeptanz seitens des Arztes und Patienten; präzisere Befunderstellung; Unterstützung des Arztes, Vermeiden von Übersehen; schnellere Befunderstellung

- **Eigene Nutzungswahrscheinlichkeit:** Auf einer Skala von 1 (= Nutzung sehr unwahrscheinlich) bis 5 (= Nutzung sehr wahrscheinlich) aus Patientensicht Nutzung unwahrscheinlich (2) aus Angst, dass sich der Arzt komplett auf die KI verlassen könnte.
- **GKV-Routinedaten für Parkinson-Sturz-Prävention:**
 - **Entwicklungsstand:** In der Entwicklung
 - **Chancen und Risiken:** Risiko einer ggf. nicht ausreichenden Datenlage zur Prädiktion; Möglichkeit eines frühzeitigen Agierens; verbessertes Verständnis der Krankheitsentstehung; Steigerung von Lebensqualität; Sensibilisierung von Patienten und Angehörigen; Kostenreduktion durch Vermeidung von Inanspruchnahme von Leistungserbringern
 - **Eigene Nutzungswahrscheinlichkeit:** Auf einer Skala von 1 (= Nutzung sehr unwahrscheinlich) bis 5 (= Nutzung sehr wahrscheinlich) aus Patientensicht Nutzung sehr wahrscheinlich (5) aufgrund der höheren Sensibilisierung von Patienten für Stürze.
- **Longitudinale Datenauswertung/Krankheitsprädiktion:**
 - **Entwicklungsstand:** In der Entwicklung
 - **Chancen und Risiken:** Risiko eines fehlerhaften Algorithmus; Frage nach der Haftung für KI-Entscheidungen auf Seiten des Arztes; Angst vor Ersatz durch KI seitens des Arztes; Kostenersparnis durch individuelle und gezielte Therapie; gezieltere Prädiktion; schnellere Diagnose; verbesserte Therapiemöglichkeit; Zeitersparnis; Unterstützung für Ärzte; Möglichkeit der Nutzbarmachung von bisher ungenutzten Datenmengen
 - **Eigene Nutzungswahrscheinlichkeit:** Auf einer Skala von 1 (= Nutzung sehr unwahrscheinlich) bis 5 (= Nutzung sehr wahrscheinlich) aus Patientensicht Nutzung sehr wahrscheinlich (5).
- **Verletzungsdiagnose paralympischer Athleten:**
 - **Entwicklungsstand:** In der Anwendung, aber kein Produkt
 - **Chancen und Risiken:** Frage nach der Haftung für KI-Entscheidungen auf Seiten des Arztes; Risiko fehlerhafter Algorithmen bzw. nicht auf den Einzelfall zutreffender Statistiken; Risiko einer zunehmenden Kommerzialisierung des Sports; Reduzierung von personellen Ausfällen und damit finanziellen Einbußen auf Seiten des Sportverbandes; verbesserte Präventionsmöglichkeit für Sportler
 - **Eigene Nutzungswahrscheinlichkeit:** Auf einer Skala von 1 (= Nutzung sehr unwahrscheinlich) bis 5 (= Nutzung sehr wahrscheinlich) aus Patientensicht Nutzung sehr wahrscheinlich (5) sowie aus Arztsicht Nutzung wahrscheinlich (4).
- **Autismustherapie:**
 - **Entwicklungsstand:** In der Entwicklung
 - **Chancen und Risiken:** Hohe Investitionskosten; Angst vor Ersatz des Therapeuten durch Roboter bzw. KI; Risiko einer emotionalen Bindung des Kindes zum Roboter; Risiko des Übernehmens der roboterhaften Emotionsdarstellung seitens des Kindes; Kostenersparnis im Gesundheitssystem durch unnötige Langzeittherapie; Zeitersparnis für Therapeuten durch verbesserte Hilfestellung des Roboters für das Kind und dadurch mehr Zeit für das Kind; verbesserte Therapie für technikaffine Kinder im Autismus-Spektrum
 - **Eigene Nutzungswahrscheinlichkeit:** Auf einer Skala von 1 (= Nutzung sehr unwahrscheinlich) bis 5 (= Nutzung sehr wahrscheinlich) aus Patienten- bzw. Angehörigensicht Nutzung sehr wahrscheinlich (5), sofern es sich um eine Form des Autismus handelt, bei der die KI-gestützte Therapie eingesetzt werden kann.

- **Blutbildanalyse Hämacam:**
 - **Entwicklungsstand:** Lizenziert und verkauft; inzwischen insolvent
 - **Chancen und Risiken:** Arbeitserleichterung und höhere Sicherheit für Labore durch Automatisierung; Zeit- und Geldgewinn für Labore; sicheres Produkt, da finale Abnahme und Entscheidung beim Menschen liegt
 - **Eigene Nutzungswahrscheinlichkeit:** Auf einer Skala von 1 (= Nutzung sehr unwahrscheinlich) bis 5 (= Nutzung sehr wahrscheinlich) sowohl aus Arzt- als auch Patientensicht Nutzung sehr wahrscheinlich (5), da Mechanismen eingebaut wurden, „die auch aus Medizinprodukte-Sicht und juristischer Sicht sichergestellt haben, dass eben der Computer nicht das letzte Wort hat.“
- **App für Kardiopatienten:**
 - **Entwicklungsstand:** App existiert als DiGa; Algorithmen in der Entwicklung
 - **Chancen und Risiken:** Furcht vor möglichem Ersatz durch die KI sowie Verlust des Kontakts zum Patienten auf Seiten des Arztes; Risiko möglicher Fehler im Algorithmus bzw. der Sensorik; Patient wiegt sich ggf. fälschlicherweise in Sicherheit und wird in Bezug auf die eigene Krankheit unvorsichtig; Zeitersparnis/Entlastung für Arzt; Entlastung der Krankenkassen durch Vermeidung kostspieliger Hospitalisierungen; verbesserte Versorgung des Patienten durch engmaschige Überwachung
 - **Eigene Nutzungswahrscheinlichkeit:** Auf einer Skala von 1 (= Nutzung sehr unwahrscheinlich) bis 5 (= Nutzung sehr wahrscheinlich) aus Arztsicht Nutzung wahrscheinlich (4) in Abhängigkeit von der Sicherheit und der notwendigen zusätzlichen Involvierung des Arztes. Aus Patientensicht Nutzung sehr wahrscheinlich (5).
- **Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten:**
 - **Entwicklungsstand:** Objektivierung der Ganganalyse als Produkt am Markt; Prädiktion zukünftiger Parkinson-Erkrankung über das Gangbild in Entwicklung
 - **Chancen und Risiken:** Notwendigkeit des Tragens des Sensorschuhs zur Diagnoseerstellung; bessere Behandlung durch objektivierbare Scores; Kosten- und Zeitersparnis für Arzt bei Beurteilung und ggf. weiteren Behandlungen
 - **Eigene Nutzungswahrscheinlichkeit:** Auf einer Skala von 1 (= Nutzung sehr unwahrscheinlich) bis 5 (= Nutzung sehr wahrscheinlich) sowohl aus Arzt- als auch aus Patientensicht Nutzung sehr wahrscheinlich (5), da „das ein bereits eingesetztes Produkt in vielen Studien, überprüfbares Produkt ist, und von daher auch von mir als entsprechend sicher eingeschätzt wird.“
- **Digitale Pathologie (medizinische Bildanalyse):**
 - **Entwicklungsstand:** Erste zugelassene Systeme am Markt
 - **Chancen und Risiken:** Notwendigkeit des Umgewöhnens des stark strukturierten Workflows bei Pathologen; hohe Investitionskosten; fehlende Serverinfrastruktur für Datenverwaltung; Frage der Haftung bei fehlerhafter KI-Entscheidung; Unterstützung des Arztes, kein Übersehen mehr; schnelleres und leichteres Einholen einer Zweitmeinung durch digitalen Versand von Bildern; Reduzierung Arbeitsaufwand durch KI-Vorsortierung; Reduzierung Kosten; Verbesserung Diagnostik, da KI Veränderungen erkennen kann, die menschliches Auge nicht erkennt; KI bietet Ausgleich für personelle Nachwuchsprobleme in der Pathologie; zukunftssträftig durch starken Push durch die Industrie durch Ausbau entsprechender Geschäftszweige (z.B. Siemens Healthineers)
 - **Eigene Nutzungswahrscheinlichkeit:** Auf einer Skala von 1 (= Nutzung sehr unwahrscheinlich) bis 5 (= Nutzung sehr wahrscheinlich) aus Arztsicht Nutzung unwahrscheinlich (2), weil es sich bei Pathologen um „eine recht konservative

Berufsgruppe, von der Altersstruktur auch eher ein bisschen älter“ handelt. Aus Sicht der Industrie hohe Nutzungswahrscheinlichkeit (4-5), da „wahnsinnig viele Unternehmen auf dieses Thema digitale Pathologie im Moment setzen.“

- **Leitlinienmodellierung/longitudinale Modellierung:**
 - **Entwicklungsstand:** In der Entwicklung
 - **Chancen und Risiken:** Möglichkeit der Unterstützung des Arztes durch Entscheidungsübernahme („da ist es einfach so komplex dass es der Mensch vielleicht einfach gar nicht mehr überblicken kann“); Risiko fehlender Akzeptanz seitens des Arztes („aber auch bei den Ärzten hast du natürlich den möglichen Nachteil dass du dem Expertentum auch so bisschen absprichst. Also wenn du ihnen halt sagst hier das KI System kann das besser entscheiden als du, oder hör darauf oder so“); möglicher Missbrauch durch Pharmaunternehmen, die bestimmtes Medikament verkaufen möchten; Möglichkeit eines ganzheitlichen Blicks auf den Patienten; vereinfachtes interdisziplinäres Arbeiten durch vereinfachte Einsichtnahme und Interpretation von Daten; Möglichkeit einer personalisierteren Behandlung für Patienten
 - **Eigene Nutzungswahrscheinlichkeit:** Auf einer Skala von 1 (= Nutzung sehr unwahrscheinlich) bis 5 (= Nutzung sehr wahrscheinlich) aus Patientensicht Nutzung sehr wahrscheinlich (5) aufgrund personalisierter Behandlungsoptionen. Aus Arztsicht Nutzung wahrscheinlich (4), da interdisziplinäres Arbeiten gewohnt.

4.2.5 Erklärung eigener KI-Beispiele für KI-Neulinge

Beim Versuch der Erklärung der eingangs genannten KI-Versorgungsbeispiele für einen KI-Neuling (z.B. Oma) griffen die TeilnehmerInnen häufig auf einen Vergleich der Fähigkeiten von KI und Arzt zurück. Hier wurden folgende Erklärungsansätze genannt:

- **Konstantes Energielevel der KI vs. abnehmendes Energielevel des Arztes:** „Der Computer ist konstant gut bei 90 Prozent. Und der Arzt wird irgendwie bei Tagesbeginn bei 100 Prozent oder bei 120 Prozent heruntergehen auf 50 Prozent.“
- **Fähigkeit zum Durchgehen großer Datenmengen bei KI vs. eingeschränkte Fähigkeit zum Durchgehen großer Datenmengen beim Arzt:** „Es gibt halt einfach wahnsinnig viele Daten und das muss alles aufbereitet werden. Und dann ihr zu sagen vielleicht dass keiner diese Vielfalt an Daten überblicken kann ist halt auch nicht unbedingt besonders vertrauenswürdig. Aber vielleicht würde ich so anfangen, du merkst ja wenn du zum Arzt gehst werden was weiß ich, super viele Untersuchungen gemacht, da wird Blut genommen, da wird das und das und das gemacht. Und der Arzt kann-, keine Ahnung du gehst zum Hausarzt, du gehst zum Kardiologen und du gehst zu dem und dem, und diese ganzen Informationen die da vorliegen, das kann sich keiner mehr durchlesen, das ist einfach viel zu viel. Und dadurch brauche ich halt vielleicht ein KI System.“
- **KI als Wissen vieler Ärzte vs. Wissen eines Arztes:** „Hey, pass auf, der Nutzen ist der, dass da das Wissen vieler Experten letzten Endes dahintersteckt und nicht nur des einen Arztes“.
- **Verbesserte Diagnose durch zweite Meinung in Form von KI:** „Und dann kann man halt sagen und es ist halt normalerweise nicht üblich, dass ein zweiter Arzt drauf guckt, weil es einfach zu teuer ist oder gar nicht so viele Ärzte vielleicht verfügbar sind. Aber dass es eigentlich manchmal schön wäre halt eine zweite Meinung noch zu haben, dass jemand zweites nochmal drauf guckt ob der erste Arzt irgendwas übersehen hat oder vielleicht falsch eingeschätzt hat. Und dann sozusagen nochmal so eine Software oder keine Ahnung, irgendeine Technik nochmal drüber guckt, die dann halt sozusagen diese zweite Meinung abbildet.“

- **Unterstützung des medizinischen Personals bei ansteigenden Patientenzahlen:** „Wir haben eine überalternde Gesellschaft. Das heißt, wir haben immer mehr Patienten und Patientinnen und immer weniger Fachärzte. Das heißt, mit immer weniger werdenden Fachärzten müssen wir die Versorgung für immer mehr Krankwerdende machen. Das heißt, wir haben pro Patient immer weniger Zeit. Um das zu kompensieren und den Arzt zu unterstützen, dass er keinen Fehler macht, dass er nicht übermüdet, dass er nichts übersieht, können Computer benutzt werden, um den Arzt, die Ärztin bei der Auswertung von Daten, beispielsweise Endoskopiebilder, Mammografiebilder, zu unterstützen und Hinweise zu geben, guck doch einmal, könnte das sein.“

Weitere, genannte Erklärungsansätze fokussieren primär darauf, mit einfachen Worten die Fähigkeiten einzelner KI-gestützter Technologien zu erläutern. Auffallend ist hierbei, dass der Fokus jeweils auf den KI-gestützten Technologien liegt, während das eigentliche Kernstück in Form der **KI nicht erwähnt** und somit nicht erklärt wird (Black Box). Basierend auf den KI-Beispielen wurden folgende Erklärungsansätze genannt:

- **Das Hemd erkennt Anomalien und holt automatisch Hilfe:** „Oma, du bist jetzt nicht mehr die Jüngste, du bist hier schon 95, du bist schon 99 und wir wissen, du lebst alleine. Und um rechtzeitig festzustellen, dass etwas bei dir ist, guck mal, hier dieser BH oder dieses Hemd, das trägst du jeden Tag. Wir bauen dir da einen Sensor hinein und der schaut, was dein Herz macht. Und wenn dein Herz nicht so macht, wie es tut oder wenn du umfallen solltest, wie letzte Woche schon einmal passiert, dann könnte der BH, das Hemd, der Sensor rechtzeitig jemanden rufen, der dir hilft.“
- **Roboter hilft Kind, Emotionen zu verstehen:** „Also dem Kind würde ich vielleicht erklären, dass der Roboter ihn unterstützt, dass er das Kind unterstützen möchte, um ihm eben zu zeigen wie man auf Emotionen reagiert, wie manche Emotionen aussehen. Und dass der Roboter eben aber auch die Emotionen des Kindes erkennen kann und darauf reagieren kann. Und im Prinzip könnte man das auch so einem KI-Neuling erzählen.“
- **Schuh misst Parkinson-Erkrankung:** „Also auch wieder relativ oberflächlich gesprochen ist das einfach ein Schuh, der misst, wie man läuft, und daraus schließen kann, ob sich die Parkinsonkrankheit verschlimmert oder verbessert hat und der Arzt da drauf hingehend da natürlich die Medikation und Behandlung besser abstimmen kann.“
- **Computer erkennt Verschlechterung und empfiehlt Arztbesuch:** „Ich glaube da würde ich auch nicht so tiefgreifend erklären, was da alles im Einzelnen passiert, da kann man sicherlich sagen, dass der Computer die App für eben feststellen kann, wenn es ihr schlechter, wenn es ihrem Herzen schlechter geht und dementsprechend frühzeitig signalisieren kann, dass sie den Arzt aufsuchen soll, um eben schlimmeres zu verhindern.“

4.2.6 Evaluation des Kategorisierungssystems hinsichtlich Vollständigkeit, Verständlichkeit und Interdisziplinarität

Tabelle 9 zeigt die Ergebnisse der abschließenden Evaluation des Kategorisierungssystems hinsichtlich der Aspekte Vollständigkeit, Verständlichkeit und Interdisziplinarität auf einer 5er-Likert-Skala mit den Endpunkten 1 (= trifft überhaupt nicht zu) und 5 (= trifft voll zu).

Tabelle 9. Einschätzung des Kategorisierungssystems

	Mittelwert	Minimum/Maximum	N
Kategorisierungssystem verständlich	4,5 (SD=,54)	4/5	6

Kategorisierungssystem vollständig	4,0 (SD=,63)	3/5	6
Kategorisierungssystem deckt wichtigste Bereiche ab	4,83 (SD=,40)	4/5	6
Kategorisierungssystem wird interdisziplinärer Sichtweise gerecht	4,33 (SD=,81)	3/5	6

Anmerkung: Angegebenen sind bei Mittelwerte inkl. des kleinsten bzw. größten Werts. SD entspricht der Standardabweichung des Mittelwertes. N gibt die Anzahl der Personen an, die die Frage beantwortet haben.

Sowohl in Bezug auf Verständlichkeit als auch Abdeckung der wichtigsten Bereiche wurde das Kategorisierungssystem von den TeilnehmerInnen zusammenfassend im Mittel als sehr gut bewertet. Die Aspekte Vollständigkeit sowie Berücksichtigung einer interdisziplinären Sichtweise wurden im Mittel als gut bewertet.

4.2.7 Finales Kategorisierungssystem

Tabelle 10. Finales Kategorisierungssystem

<i>Kategorien</i>					
Input	Versorgung	Fachgebiet	Nutzer	KI-basierte Technologie	Output
Bilddaten	Prävention	Chirurgie	Patienten	Bilderkennung und -verstehen	Umfassendere Betreuung
Audiodaten	Diagnose	Dermatologie	Ärzte	Big Data-Analyse	Verbesserte Beratung
Anamnesedaten	Therapie	Innere Medizin	Angehörige von Patienten	Mensch-Maschine-Interaktion und Assistenzsysteme	Präzisere (Ausschluss-) Diagnose
Routinedaten der GKV (Abrechnungsdaten)	Rehabilitation	Intensivmedizin	Medizinisches/-technisches Fachpersonal	Robotik und autonome Systeme	Verbesserte Detektion von Anomalien, Auffälligkeiten
Vitaldaten (diskrete Messung)	Pflege	Neurologie	Pflegerisches Fachpersonal	Biosignal- / Vitalwert- / Sensorikanalyse	Frühzeitige Intervention (Event-/Krankheitsvermeidung)
Zeitreihendaten (kontinuierliche Messung)		Neurologie	Forschende	Sprach- und Textverstehen	Verbesserte, personalisierte Behandlung
Daten aus Untersuchung, Labor, Demografie		Notarztwesen	Personal der Krankenkassen & Versicherungen	Virtuelle und augmented Realität	Verbesserte (Behandlungs-) Abläufe
Sonst. klinische Daten (handschr. Notizen, Begegnung)		Öffentliche Gesundheit	Personal der Krankenhausverwaltung	X-plainable AI	Verbesserte (Versorgungs-) Forschung & Entwicklung
"-omics" Daten		Pathologie	Personal sonstiger stationärer Einrichtungen		Verbesserte Dokumentation
Wissenschaftliche Datenlage (Fachliteratur, -studien, experten)		Pflege	Apotheker		
		Pharmakologie	Heilmittelerbringer (Physio-, Ergotherapeuten, Logopäden etc.)		
		Psychiatrie	Hilfsmittelverordner		
		Psychologie	Personal des Rettungsdienstes		
		Radiologie			
		Sportmedizin			
		Urologie			
		Onkologie			

Ausprägungen

5. Online-Fragebogen zur Stimulusauswahl auf Basis des Kategorisierungsinstruments

5.1 Methodik

Um die Vielzahl der im Rahmen der Experteninterviews generierten Beispiele für die Akzeptanzstudie zu klassifizieren und nutzbar zu machen, wurde im Anschluss an die Auswertung der Experteninterviews ein Online-Fragebogen aufgesetzt, der folgende Bereiche bzw. Fragen adressierte:

- Bewertung der KI-Beispiele aus den Versorgungsbereichen Prävention, Diagnostik und Therapie hinsichtlich Verständlichkeit, Relevanz und Entlastungsfunktion auf einer 5er-Likert-Skala
- Auswahl und Priorisierung von je drei KI-Beispielen aus den Versorgungsbereichen Prävention, Diagnostik und Therapie, die sich nach eigener Einschätzung am besten dazu eignen, um sie gleichzeitig mit einer großen Anzahl an Versicherten und Ärztinnen bezüglich KI-Akzeptanz zu testen
- Auswahl des Nutzungsszenarios bezogen auf die jeweils ausgewählten KI-Beispiele aus Arztsicht (keine Anwendung/Anwendung, wenn KI nur beobachtet oder fühlt, aber keine Entscheidung trifft (human in the loop)/Anwendung, wenn KI interpretiert oder evaluiert und Vorschläge macht (human on the loop)/Anwendung, wenn KI selbst entscheidet oder handelt (human out of the loop))
- Auswahl des Nutzungsszenarios bezogen auf die jeweils ausgewählten KI-Beispiele aus Patientensicht (keine Anwendung/Anwendung, wenn KI nur beobachtet oder fühlt, aber keine Entscheidung trifft (human in the loop)/Anwendung, wenn KI interpretiert oder evaluiert und Vorschläge macht (human on the loop)/Anwendung, wenn KI selbst entscheidet oder handelt (human out of the loop))
- Nennung der gewünschten zusätzlichen Informationen, die aus Arzt- bzw. Patientensicht nötig wären, um der KI in den gewählten Beispielen die alleinige Entscheidung bzw. Handlung zu überlassen.

5.2 Ergebnis

Im Zeitraum vom 02.-15.02.22 nahmen insgesamt 5 ExpertInnen an der Online-Befragung teil. Bei den TeilnehmerInnen handelte es sich um die ExpertInnen, die bereits an den Interviews teilgenommen hatten. Zusätzlich wurden zwei Experten der AOK Bayern angefragt, aufgrund der Anonymität der Online-Befragung kann jedoch keine Aussage darüber getroffen werden, welche der ExpertInnen tatsächlich an der Online-Befragung teilgenommen haben.

5.2.1 Bewertung der KI-Beispiele hinsichtlich Verständlichkeit, Relevanz und Entlastungsfunktion

5.2.1.1 Bewertung der KI-Beispiele – Prävention

Tabelle 11 zeigt, wie die KI-Beispiele aus dem Versorgungsbereich Prävention von den TeilnehmerInnen hinsichtlich **Verständlichkeit** für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) bewertet wurden (1 = sehr geringe Verständlichkeit, 5 = sehr hohe Verständlichkeit).

Tabelle 11. Einschätzung der KI-Beispiele aus der Prävention bzgl. Verständlichkeit

	Mittelwert (SD)	Min/Max	N
P1: KI-basierte Detektion von Läsionen/Polypen	4,80 (SD=,44)	4/5	5
P2: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	3,80 (SD=,44)	3/4	5
P3: KI-basierte Erkennung von Brustkrebs	4,60 (SD=,54)	4/5	5
P4: GKV-Routinedaten für Parkinson-Sturzprävention	3,20 (SD=1,78)	1/5	5
P5: Verletzungsdiagnose paralympischer Athleten	4,00 (SD=1,00)	3/5	5

Anmerkung: Angegebenen sind bei Mittelwerte inkl. des kleinsten bzw. größten Werts. SD entspricht der Standardabweichung des Mittelwertes. N gibt die Anzahl der Personen an, die die Frage beantwortet haben.

Tabelle 12 zeigt, wie die KI-Beispiele aus dem Versorgungsbereich Prävention von den TeilnehmerInnen hinsichtlich **Relevanz** für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) bewertet wurden (1 = sehr geringe Relevanz, 5 = sehr hohe Relevanz).

Tabelle 12. Einschätzung der KI-Beispiele aus der Prävention bzgl. Relevanz

	Mittelwert (SD)	Min/Max	N
P1: KI-basierte Detektion von Läsionen/Polypen	4,60 (SD=,54)	4/5	5
P2: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	3,00 (SD=1,00)	2/4	5
P3: KI-basierte Erkennung von Brustkrebs	4,60 (SD=,54)	4/5	5
P4: GKV-Routinedaten für Parkinson-Sturzprävention	4,20 (SD=,83)	3/5	5
P5: Verletzungsdiagnose paralympischer Athleten	3,40 (SD=,89)	3/5	5

Anmerkung: Angegebenen sind bei Mittelwerte inkl. des kleinsten bzw. größten Werts. SD entspricht der Standardabweichung des Mittelwertes. N gibt die Anzahl der Personen an, die die Frage beantwortet haben.

Tabelle 13 zeigt, wie die KI-Beispiele aus dem Versorgungsbereich Prävention von den TeilnehmerInnen hinsichtlich ihrer **Entlastungsfunktion** für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) bewertet wurden (1 = sehr geringe Entlastung, 5 = sehr hohe Entlastung).

Tabelle 13. Einschätzung der KI-Beispiele aus der Prävention bzgl. Entlastungsfunktion

	Mittelwert (SD)	Min/Max	N
P1: KI-basierte Detektion von Läsionen/Polypen	4,20 (SD=,83)	3/5	5
P2: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	2,80 (SD=1,30)	2/5	5
P3: KI-basierte Erkennung von Brustkrebs	4,20 (SD=,83)	3/5	5
P4: GKV-Routinedaten für Parkinson-Sturzprävention	3,80 (SD=,83)	3/5	5
P5: Verletzungsdiagnose paralympischer Athleten	2,80 (SD=1,48)	1/5	5

Anmerkung: Angegebenen sind bei Mittelwerte inkl. des kleinsten bzw. größten Werts. SD entspricht der Standardabweichung des Mittelwertes. N gibt die Anzahl der Personen an, die die Frage beantwortet haben.

5.2.1.2 Bewertung der KI-Beispiele – Diagnostik

Tabelle 14 zeigt, wie die KI-Beispiele aus dem Versorgungsbereich Diagnostik von den TeilnehmerInnen hinsichtlich **Verständlichkeit** für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) bewertet wurden (1 = sehr geringe Verständlichkeit, 5 = sehr hohe Verständlichkeit).

Tabelle 14. Einschätzung der KI-Beispiele aus der Diagnostik bzgl. Verständlichkeit

	Mittelwert (SD)	Min/Max	N
D1: Multimodale Zeitreihendaten (Fitnessshirt-Monitoring)	3,20 (SD=1,64)	1/5	5
D2: KI-basierte Schmerzanalyse in der Pflege	3,80 (SD=,83)	3/5	5
D3: Ultraschallbasierte Flüssigkeitsgabe im Notarztwesen	3,25 (SD=1,70)	1/5	4
D4: Blutbildanalyse Hämacam	3,60 (SD=1,14)	2/5	5
D5: App für Kardiopatienten	4,00 (SD=,70)	3/5	5
D6: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	3,40 (SD=,89)	2/4	5
D7: Digitale Pathologie (medizin. Bildanalyse)	3,40 (SD=,89)	2/4	5
D8: Leitlinienmodellierung/longitudinale Modellierung	4,00 (SD=1,00)	3/5	5

Anmerkung: Angegebenen sind bei Mittelwerte inkl. des kleinsten bzw. größten Werts. SD entspricht der Standardabweichung des Mittelwertes. N gibt die Anzahl der Personen an, die die Frage beantwortet haben.

Tabelle 15 zeigt, wie die KI-Beispiele aus dem Versorgungsbereich Diagnostik von den TeilnehmerInnen hinsichtlich **Relevanz** für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) bewertet wurden (1 = sehr geringe Relevanz, 5 = sehr hohe Relevanz).

Tabelle 15. Einschätzung der KI-Beispiele aus der Diagnostik bzgl. Relevanz

	Mittelwert (SD)	Min/Max	N
D1: Multimodale Zeitreihendaten (Fitnessshirt-Monitoring)	3,00 (SD=1,22)	1/4	5
D2: KI-basierte Schmerzanalyse in der Pflege	4,00 (SD=1,22)	2/5	5
D3: Ultraschallbasierte Flüssigkeitsgabe im Notarztwesen	3,25 (SD=1,70)	1/5	4
D4: Blutbildanalyse Hämacam	4,00 (SD=,70)	3/5	5
D5: App für Kardiopatienten	3,80 (SD=,83)	3/5	5
D6: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	3,50 (SD=,57)	3/4	4

D7: Digitale Pathologie (medizin. Bildanalyse)	4,60 (SD=,54)	4/5	5
D8: Leitlinienmodellierung/longitudinale Modellierung	4,00 (SD=1,00)	3/5	5

Anmerkung: Angegebenen sind bei Mittelwerte inkl. des kleinsten bzw. größten Werts. SD entspricht der Standardabweichung des Mittelwertes. N gibt die Anzahl der Personen an, die die Frage beantwortet haben.

Tabelle 16 zeigt, wie die KI-Beispiele aus dem Versorgungsbereich Diagnostik von den TeilnehmerInnen hinsichtlich ihrer **Entlastungsfunktion** für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) bewertet wurden (1 = sehr geringe Entlastung, 5 = sehr hohe Entlastung).

Tabelle 16. Einschätzung der KI-Beispiele aus der Diagnostik bzgl. Entlastungsfunktion

	Mittelwert (SD)	Min/Max	N
D1: Multimodale Zeitreihendaten (Fitnessshirt-Monitoring)	2,60 (SD=1,14)	1/4	5
D2: KI-basierte Schmerzanalyse in der Pflege	3,40 (SD=1,67)	1/5	5
D3: Ultraschallbasierte Flüssigkeitsgabe im Notarztwesen	3,33 (SD=2,08)	1/5	3
D4: Blutbildanalyse Hämacam	3,80 (SD=,44)	3/4	5
D5: App für Kardiopatienten	3,80 (SD=1,30)	2/5	5
D6: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	3,20 (SD=1,48)	1/5	5
D7: Digitale Pathologie (medizin. Bildanalyse)	4,40 (SD=,54)	4/5	5
D8: Leitlinienmodellierung/longitudinale Modellierung	3,40 (SD=,89)	3/5	5

Anmerkung: Angegebenen sind bei Mittelwerte inkl. des kleinsten bzw. größten Werts. SD entspricht der Standardabweichung des Mittelwertes. N gibt die Anzahl der Personen an, die die Frage beantwortet haben.

5.2.1.3 Bewertung der KI-Beispiele – Therapie

Tabelle 17 zeigt, wie die KI-Beispiele aus dem Versorgungsbereich Therapie von den TeilnehmerInnen hinsichtlich **Verständlichkeit** für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) bewertet wurden (1 = sehr geringe Verständlichkeit, 5 = sehr hohe Verständlichkeit).

Tabelle 17. Einschätzung der KI-Beispiele aus der Therapie bzgl. Verständlichkeit

	Mittelwert (SD)	Min/Max	N
T1: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	3,50 (SD=,577)	3/4	4
T2: Genetisch spezifizierte Antibiotikagabe im Krankenhaus	3,80 (SD=1,09)	2/5	5
T3: Ultraschallbasierte Flüssigkeitsgabe im Notarztwesen	3,25 (SD=,95)	2/4	4
T4: Frühzeitige Erkennung von Unterzuckerung auf der Intensivstation	3,80 (SD=,83)	3/5	5
T5: Frühzeitige Erkennung von Unterzuckerung für den Privatgebrauch	3,60 (SD=1,14)	2/5	5
T6: Longitudinale Datenauswertung/Krankheitsprädiktion	3,00 (SD=1,00)	2/4	5
T7: Autismustherapie	3,20 (SD=,83)	2/4	5
T8: App für Kardiopatienten	3,75 (SD=,50)	3/4	4
T9: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	3,50 (SD=1,00)	2/4	4
T10: Leitlinienmodellierung/Longitudinale Modellierung	3,25 (SD=,50)	3/4	4

Anmerkung: Angegebenen sind bei Mittelwerte inkl. des kleinsten bzw. größten Werts. SD entspricht der Standardabweichung des Mittelwertes. N gibt die Anzahl der Personen an, die die Frage beantwortet haben.

Tabelle 18 zeigt, wie die KI-Beispiele aus dem Versorgungsbereich Therapie von den TeilnehmerInnen hinsichtlich **Relevanz** für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) bewertet wurden (1 = sehr geringe Relevanz, 5 = sehr hohe Relevanz).

Tabelle 18. Einschätzung der KI-Beispiele aus der Therapie bzgl. Relevanz

	Mittelwert (SD)	Min/Max	N
T1: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	2,75 (SD=,95)	2/4	4
T2: Genetisch spezifizierte Antibiotikagabe im Krankenhaus	4,00 (SD=,70)	3/5	5
T3: Ultraschallbasierte Flüssigkeitsgabe im Notarztwesen	3,00 (SD=1,41)	1/4	4
T4: Frühzeitige Erkennung von Unterzuckerung auf der Intensivstation	4,00 (SD=,70)	3/5	5
T5: Frühzeitige Erkennung von Unterzuckerung für den Privatgebrauch	4,00 (SD=,70)	3/5	5
T6: Longitudinale Datenauswertung/Krankheitsprädiktion	3,50 (SD=,57)	3/4	4
T7: Autismustherapie	3,40 (SD=,54)	3/4	5
T8: App für Kardiopatienten	4,25 (SD=,50)	4/5	4
T9: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	4,00 (SD=,81)	3/5	4
T10: Leitlinienmodellierung/Longitudinale Modellierung	3,75 (SD=,50)	3/4	4

Anmerkung: Angegebenen sind bei Mittelwerte inkl. des kleinsten bzw. größten Werts. SD entspricht der Standardabweichung des Mittelwertes. N gibt die Anzahl der Personen an, die die Frage beantwortet haben.

Tabelle 19 zeigt, wie die KI-Beispiele aus dem Versorgungsbereich Therapie von den TeilnehmerInnen hinsichtlich ihrer **Entlastungsfunktion** für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) bewertet wurden (1 = sehr geringe Entlastung, 5 = sehr hohe Entlastung).

Tabelle 19. Einschätzung der KI-Beispiele aus der Therapie bzgl. Entlastungsfunktion

	Mittelwert (SD)	Min/Max	N
T1: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	3,25 (SD=1,50)	2/5	4
T2: Genetisch spezifizierte Antibiotikagabe im Krankenhaus	3,40 (SD=,89)	2/4	5
T3: Ultraschallbasierte Flüssigkeitsgabe im Notarztwesen	2,75 (SD=,95)	2/4	4
T4: Frühzeitige Erkennung von Unterzuckerung auf der Intensivstation	4,20 (SD=,44)	4/5	5
T5: Frühzeitige Erkennung von Unterzuckerung für den Privatgebrauch	4,20 (SD=,44)	4/5	5
T6: Longitudinale Datenauswertung/Krankheitsprädiktion	3,50 (SD=,57)	3/4	4
T7: Autismustherapie	4,00 (SD=,70)	3/5	5
T8: App für Kardiopatienten	3,67 (SD=,57)	3/4	3
T9: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	3,75 (SD=,95)	3/5	4
T10: Leitlinienmodellierung/Longitudinale Modellierung	2,75 (SD=,95)	2/4	4

Anmerkung: Angegebenen sind bei Mittelwerte inkl. des kleinsten bzw. größten Werts. SD entspricht der Standardabweichung des Mittelwertes. N gibt die Anzahl der Personen an, die die Frage beantwortet haben.

5.2.2 Auswahl und Priorisierung von KI-Beispielen für KI-Akzeptanzstudie

5.2.2.1 Auswahl und Priorisierung von KI-Beispielen – Prävention

Tabelle 20 zeigt die Auswahl der von den TeilnehmerInnen als **am besten für die KI-Akzeptanzstudie geeigneten KI-Beispiele** aus dem Versorgungsbereich Prävention.

Tabelle 20. Auswahl der am besten geeigneten KI-Beispiele aus der Prävention

	Häufigkeit	Prozent	N
Rang 1: Am besten geeignetes KI-Beispiel aus der Prävention für gleichzeitige Akzeptanz-Testung mit Versicherten und ÄrztInnen			
P1: KI-basierte Detektion von Läsionen/Polypen	3	60,0	5
P3: KI-basierte Erkennung von Brustkrebs	2	40,0	5
Rang 2: Am zweitbesten geeignetes KI-Beispiel aus der Prävention für gleichzeitige Akzeptanz-Testung mit Versicherten und ÄrztInnen			
P1: KI-basierte Detektion von Läsionen/Polypen	1	20,0	5
P2: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	2	40,0	5
P3: KI-basierte Erkennung von Brustkrebs	2	40,0	5
Rang 3: Am drittbesten geeignetes KI-Beispiel aus der Prävention für gleichzeitige Akzeptanz-Testung mit Versicherten und ÄrztInnen			
P1: KI-basierte Detektion von Läsionen/Polypen	1	20,0	5
P3: KI-basierte Erkennung von Brustkrebs	1	20,0	5
P4: GKV-Routinedaten zur Parkinson-Sturz-Prävention	1	20,0	5
P5: Verletzungsdiagnose paralympischer Athleten	2	40,0	5

Als **Gründe** für die Auswahl der jeweiligen Beispiele wurden von den TeilnehmerInnen folgende Aspekte angeführt:

- **P1: KI-basierte Erkennung von Brustkrebs:**
 - „Bei der Erkennung von Brustkrebs wird der Vorteil direkt deutlich und es ist einfach verständlich (z.B. durch geringe Komplexität der Input-Daten).“
 - „Brustkrebs betrifft ca. 10% aller Frauen.“
 - „KI auf Basis von Bildern/Sprache eingängiger als auf Basis von AR und Versichertendaten.“
 - „Die Medizin arbeitet schon mit KI und so werden die Anwendungen im besten Fall schneller integriert. Der „Aufwand“ der Mediziner ist gering, da sich an der Routine wenig ändert, kein „Umlernen“ notwendig, die Relevanz wird schneller erkannt.“
- **P3: KI-basierte Detektion von Läsionen/Polypen:**

- „Polypendetektion ähnlich wie bei Brustkrebs und zudem gibt es einen Alltagsbezug (die meisten Leute kennen jemanden/haben selbst eine Darmspiegelung hinter sich).“
- „Koloraktale Karzinome ist Nummer 2 aller Krebserkrankungen weltweit bei Männern und Frauen.“
- „KI auf Basis von Bildern/Sprache eingängiger als auf Basis von AR und Versicherungendaten.“
- „Die Medizin arbeitet schon mit KI und so werden die Anwendungen im besten Fall schneller integriert. Der „Aufwand“ der Mediziner ist gering, da sich an der Routine wenig ändert, kein „Umlernen“ notwendig, die Relevanz wird schneller erkannt.“
- **P2: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses:**
 - „Bei der Unterstützung der Morgenprozesse ist das Problem (Überlastung des Pflegepersonals) einfach zu verstehen und jeder hat einen Alltags-Bezug (kennt jemanden im Pflegeheim), allerdings ist fraglich, ob die KI-basierte Lösung das Problem auch wirklich adressiert – hier ist Feedback sehr interessant.“
 - „In der Pflege sind arbeits- und auswertungsunterstützende Systeme dringend notwendig. Zudem erleichtert das System die Doku (nicht nur im Hinblick auf nicht muttersprachliche MitarbeiterInnen). Fehlerzahl (Zahlendreher etc.) wird sich verringern. Die Diagnostik wird (im besten Fall) schneller und sicherer.“
- **P4: GKV-Routinedaten zur Parkinson-Sturz-Prävention:**
 - „Parkinson betrifft 10-15% aller Senioren. Tendenz steigend. Die Anzahl paralympischer Athleten ist dagegen marginal. AR-Brillen leiden unter Akzeptanzproblemen und lösen meiner Meinung nach nicht das Problem. Ein Beamer für die Projektion der Daten würde hier m.E. vollkommen ausreichen.“

5.2.2.2 Auswahl und Priorisierung von KI-Beispielen – Diagnostik

Tabelle 21 zeigt die Auswahl der von den TeilnehmerInnen als **am besten für die KI-Akzeptanzstudie geeigneten KI-Beispiele** aus dem Versorgungsbereich Diagnostik.

Tabelle 21. Auswahl der am besten geeigneten KI-Beispiele aus der Diagnostik

	Häufigkeit	Prozent	N
Rang 1: Am besten geeignetes KI-Beispiel aus der Diagnostik für gleichzeitige Akzeptanz-Testung mit Versicherten und ÄrztInnen			
D1: Multimodale Zeitreihendaten (Fitnessshirt-Monitoring)	1	20,0	5
D2: KI-basierte Schmerzanalyse in der Pflege	2	40,0	5
D5: App für Kardiopatienten	1	20,0	5
D7: Digitale Pathologie (medizinische Bildanalyse)	1	20,0	5
Rang 2: Am zweitbesten geeignetes KI-Beispiel aus der Diagnostik für gleichzeitige Akzeptanz-Testung mit Versicherten und ÄrztInnen			
D2: KI-basierte Schmerzanalyse in der Pflege	2	40,0	5
D5: App für Kardiopatienten	2	40,0	5
D7: Digitale Pathologie (medizinische Bildanalyse)	1	20,0	5
Rang 3: Am drittbesten geeignetes KI-Beispiel aus der Diagnostik für gleichzeitige Akzeptanz-Testung mit Versicherten und ÄrztInnen			
D1: Multimodale Zeitreihendaten (Fitnessshirt-Monitoring)	1	20,0	5
D4: Blutbildanalyse Hämacam	1	20,0	5
D5: App für Kardiopatienten	1	20,0	5
D6: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	1	20,0	5
D8: Leitlinienmodellierung/longitudinale Modellierung	1	20,0	5

Als **Gründe** für die Auswahl der jeweiligen Beispiele wurden von den TeilnehmerInnen folgende Aspekte angeführt:

- **D1: Multimodale Zeitreihendaten (Fitnessshirt-Monitoring):**
 - „Die Daten sind allen Zielgruppen „bekannt“ und können größtenteils interpretiert werden. Die regelmäßig und in gewissen Abständen dauerhafte „Abfrage“ kann Anomalien schnell verdeutlichen.“
 - „Herz-Kreislauf-Erkrankungen sind aktuell weltweit eine der schlimmsten Risiken (zu wenig Bewegung, Übergewicht, falsche Ernährung).“
- **D2: KI-basierte Schmerzanalyse in der Pflege:**

- „Der Bedarf an Systemen, „Ereignisse“ im Körper (dazu gehört vor allem der Schmerz) bei Personen mit eingeschränkter bis nicht mehr vorhandener Artikulationsmöglichkeit zu erkennen ist hoch. Es würde die Leiden der Betroffenen deutlichst mindern. Dies mit allen damit verbundenen Folgen.“
- „Schmerzanalyse in der Pflege ist verständlich und alltagsnah, der potenzielle Vorteil ist direkt verständlich und die Antworten, ob das wirklich nützlich ist/akzeptiert wird, ist sehr interessant.“
- **D5: App für Kardiopatienten:**
 - „Kardio-App ist ebenfalls verständlich und alltagsnah und der Nutzen wird klar.“
- **D6: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten:**
 - „Parkinson betrifft weltweit ca. 10-15% aller Menschen, das muss rechtzeitig erkannt und behandelt werden.“
- **D7: Digitale Pathologie (medizinische Bildanalyse):**
 - „Digitale Pathologie (inkl. Hämatologie) ist der diagnostische Endpunkt fast aller Erkrankungen und muss dringend unterstützt werden.“
- **D8: Leitlinienmodellierung/longitudinale Modellierung:**
 - „Leitlinienmodellierung ist super interessant und ein großes Thema, allerdings muss man das v.a. für Patienten verständlicher erklären.“

5.2.2.3 Auswahl und Priorisierung von KI-Beispielen – Therapie

Tabelle 22 zeigt die Auswahl der von den TeilnehmerInnen als **am besten für die KI-Akzeptanzstudie geeigneten KI-Beispiele** aus dem Versorgungsbereich Therapie.

Tabelle 22. Auswahl der am besten geeigneten KI-Beispiele aus der Therapie

	Häufigkeit	Prozent	N
Rang 1: Am besten geeignetes KI-Beispiel aus der Therapie für gleichzeitige Akzeptanz-Testung mit Versicherten und ÄrztInnen			
T1: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	1	20,0	5
T2: Genetisch spezifizierte Antibiotikagabe im Krankenhaus	1	20,0	5
T4: Frühzeitige Erkennung von Unterzuckerung auf der Intensivstation	1	20,0	5
T5: Frühzeitige Erkennung von Unterzuckerung für den Privatgebrauch	1	20,0	5
T8: App für Kardiopatienten	1	20,0	5
Rang 2: Am zweitbesten geeignetes KI-Beispiel aus der Therapie für gleichzeitige Akzeptanz-Testung mit Versicherten und ÄrztInnen			
T4: Frühzeitige Erkennung von Unterzuckerung auf der Intensivstation	2	40,0	5
T5: Frühzeitige Erkennung von Unterzuckerung für den Privatgebrauch	2	40,0	5
T8: App für Kardiopatienten	1	20,0	5
Rang 3: Am drittbesten geeignetes KI-Beispiel aus der Therapie für gleichzeitige Akzeptanz-Testung mit Versicherten und ÄrztInnen			
T2: Genetisch spezifizierte Antibiotikagabe im Krankenhaus	2	40,0	5
T8: App für Kardiopatienten	2	40,0	5
T9: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	1	20,00	5

Als **Gründe** für die Auswahl der jeweiligen Beispiele wurden von den TeilnehmerInnen folgende Aspekte angeführt:

- **T1: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses:**
 - „Personenentlastendes Patientenmonitoring in kritischen Situationen, gerade zu Zeiten von Fachkräftemangel in der Pflege.“
- **T2: Genetisch spezifizierte Antibiotikagabe im Krankenhaus:**
 - „Unterzuckerung (privat und im KH) ist alltagsnah, verständlich und die Relevanz wird auf den ersten Blick klar genauso wie bei der Antibiotikagabe. Hier ist der Nutzen klar, allerdings würde es den Ablauf der Behandlung verändern (Laboruntersuchung bevor man das Medikament gibt), daher ist die Frage nach der Akzeptanz bei Ärzten interessant.“

- **T4: Frühzeitige Erkennung von Unterzuckerung auf der Intensivstation:**
 - „Unterzuckerung (privat und im KH) ist alltagsnah, verständlich und die Relevanz wird auf den ersten Blick klar genauso wie bei der Antibiotikagabe.“
- **T5: Frühzeitige Erkennung von Unterzuckerung für den Privatgebrauch:**
 - „Unterzuckerung (privat und im KH) ist alltagsnah, verständlich und die Relevanz wird auf den ersten Blick klar genauso wie bei der Antibiotikagabe.“
- **T8: App für Kardiopatienten:**
 - „Herz-Kreislauf-Erkrankungen müssen rechtzeitig erkannt und behandelt werden.“
- **T9: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten:**
 - „Einfache Anwendung in der Praxis im Vergleich zu allen anderen Möglichkeiten. Bei Parkinson evtl. Problem der Probandenzahl. Evtl. Erweiterung auf andere Krankheiten mit Gangstörung, z.B. MS?“

Bei der Auswahl der am besten geeigneten KI-Beispiele für die Akzeptanz-Studie fällt in allen drei Versorgungsbereichen auf, dass sich neben Begründungen im Sinne von „Krankheit betrifft X Prozent aller Personen, ist leicht verständlich“ auch Begründungen einer Notwendigkeit des Einsatzes einer bestimmten KI finden. Da mit einer Notwendigkeit eines KI-Einsatzes (z.B. aufgrund von Personalmangel, besserer Behandlungsqualität) jedoch nicht notwendigerweise eine bessere Eignung für die Zielgruppen der Akzeptanzstudie verbunden ist, sind die von den TeilnehmerInnen gewählten KI-Beispiele nochmals kritisch zu reflektieren.

5.2.3 Auswahl des Nutzungsszenarios aus Arzt- bzw. Patientensicht

Versorgungsbereich Prävention

Bei der Auswahl des Nutzungsszenarios bezogen auf die jeweils ausgewählten KI-Beispiele aus Patientensicht ergab sich ein deutliches Bild: Mit Ausnahme einer Person gaben alle TeilnehmerInnen an, dass sie alle von ihnen gewählten KI-Beispiele als Patient an sich anwenden lassen würden, wenn die KI interpretiert oder evaluiert und Vorschläge macht (human on the loop). Eine Person konnte sich vorstellen, der KI in den gewählten Beispielen aus Patientensicht auch die alleinige Entscheidung zu überlassen, wenn zusätzliche Informationen wie Hintergrunddaten, Anwendungszeitraum, Erfahrungen und Kompetenz der Menschen beim Bedienen zur Verfügung stünden. 4 Personen gaben an, der KI in den gewählten Beispielen auch mit zusätzlichen Informationen nicht die alleinige Entscheidung überlassen zu wollen.

Aus Arztsicht ergab sich bezüglich der Auswahl des Nutzungsszenarios bezogen auf die jeweils ausgewählten KI-Beispiele ein leicht differenzierteres Bild: Hier gaben alle Probanden an, der KI in den gewählten Beispielen nicht die alleinige Entscheidung überlassen zu wollen, unabhängig von weiteren zur Verfügung gestellten Informationen. Aus Arztsicht würde die Mehrheit der Probanden die gewählten KI-Beispiele verwenden, wenn die KI interpretiert oder evaluiert und Vorschläge macht (human on the loop). Bezogen auf jeweils ein KI-Fallbeispiel würde eine Person das KI-Beispiel nicht selbst verwenden bzw. das KI-Beispiel nur verwenden, wenn die KI nur beobachtet oder fühlt, aber keine Entscheidung trifft (human in the loop). Abbildung 6 zeigt die Ergebnisse für die ausgewählten KI-Beispiele aus dem Bereich Prävention:

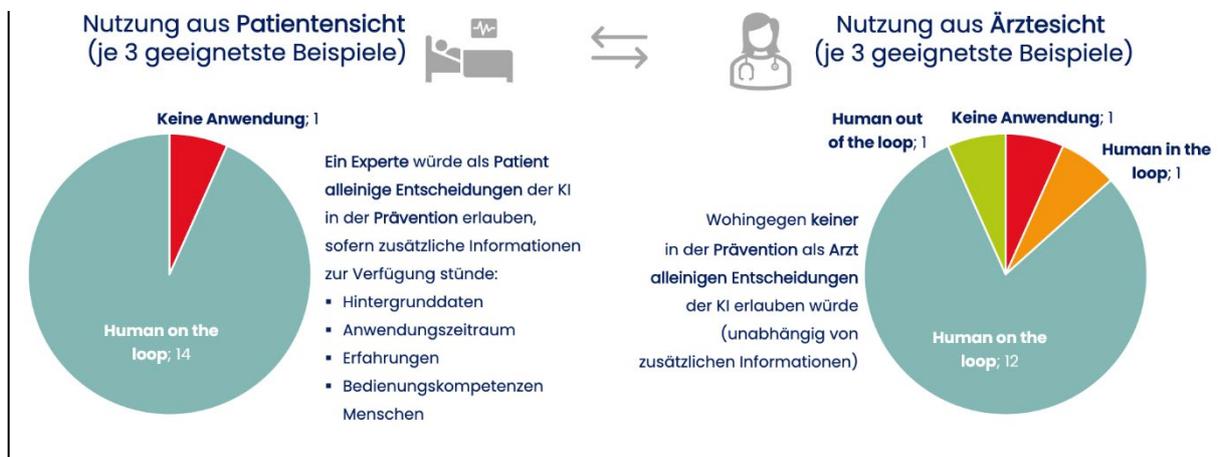


Abbildung 6. Ergebnisse für die ausgewählten KI-Beispiele - Bereich Prävention

Versorgungsbereich Diagnostik

Im Vergleich zum Versorgungsbereich Prävention zeigten sich sowohl aus Arzt- als auch aus Patientensicht ähnliche Ergebnisse im Hinblick auf die Auswahl des Nutzungsszenarios: Aus Arztsicht gaben alle Probanden an, der KI in den gewählten Beispielen nicht die alleinige Entscheidung überlassen zu wollen, unabhängig von weiteren zur Verfügung gestellten Informationen. Die Mehrheit würde jedoch auch im Versorgungsbereich Diagnostik die gewählten KI-Beispiele selbst verwenden, wenn die KI interpretiert oder evaluiert und Vorschläge macht (human on the loop). Bezogen auf jeweils ein KI-Fallbeispiel würde je eine Person das KI-Beispiel nicht selbst verwenden.

Aus Patientensicht gaben ebenfalls alle Probanden an, der KI in den gewählten Beispielen unabhängig von weiteren zur Verfügung gestellten Informationen nicht die alleinige Entscheidung überlassen zu wollen. Die Mehrheit würde jedoch die gewählten KI-Beispiele an sich anwenden lassen, wenn die KI interpretiert oder evaluiert und Vorschläge macht (human on the loop). Bezogen auf jeweils ein KI-Fallbeispiel würde je eine Person die KI nicht an sich anwenden lassen bzw. nur an sich anwenden lassen, wenn die KI nur beobachtet oder fühlt, aber keine Entscheidung trifft (human in the loop). Ein Proband gab an, die KI auch an sich anwenden zu lassen, wenn die KI selbst entscheidet oder handelt (human out of the loop). Abbildung 7 zeigt die Ergebnisse für die ausgewählten KI-Beispiele aus dem Bereich Diagnostik:



Abbildung 7. Ergebnisse für die ausgewählten KI-Beispiele - Bereich Diagnostik

Versorgungsbereich Therapie

Die Ergebnisse im Versorgungsbereich Therapie decken sich mit denen der Bereiche Prävention und Diagnostik: Aus Arztsicht gaben alle Probanden an, der KI in den gewählten Beispielen nicht die alleinige Entscheidung überlassen zu wollen, unabhängig von weiteren zur Verfügung gestellten Informationen. Die Mehrheit würde jedoch die gewählten KI-Beispiele selbst verwenden, wenn die KI interpretiert oder evaluiert und Vorschläge macht (human on the loop).

Aus Patientensicht gaben alle Probanden an, der KI in den gewählten Beispielen nicht die alleinige Entscheidung zu überlassen. Auch hier gab die Mehrheit an, die KI im gewählten Beispiel an sich anwenden zu lassen, wenn die KI interpretiert oder evaluiert und Vorschläge macht (human on the loop). Ein Proband gab an, die KI auch an sich anwenden zu lassen, wenn die KI selbst entscheidet oder handelt (human out of the loop).

Zusammenfassend lässt sich für die Versorgungsbereiche Prävention, Diagnostik und Therapie festhalten, dass der KI in den gewählten Beispielen sowohl aus Arzt- als auch Patientensicht eine gewisse Interpretations- bzw. Vorschlagskompetenz zugesprochen wird, die Entscheidungsgewalt jedoch vornehmlich beim Menschen gesehen wird. Interessant ist hier insb. die Patientensicht, bei der in allen drei Versorgungsbereichen jeweils ein Proband angab, die KI an sich anwenden zu lassen, wenn diese selbst entscheidet oder handelt (human out of the loop). Gleichzeitig wurde in zwei von drei Fällen angegeben, dass man der KI auch bei Vorlage zusätzlicher Informationen nicht die alleinige Entscheidung überlassen wolle. Dies impliziert, dass eine differenziertere Betrachtung zwischen der Entscheidungsfähigkeit der KI (mit eventueller Korrekturmöglichkeit durch den Menschen) und der alleinigen Entscheidungsfähigkeit der KI nötig ist, da diese mit unterschiedlicher Akzeptanz verbunden zu sein scheint. Abbildung 8 zeigt die Ergebnisse für die ausgewählten KI-Beispiele aus dem Bereich Therapie:



Abbildung 8. Ergebnisse für die ausgewählten KI-Beispiele - Bereich Therapie

5.2.4 Finale Auswahl der KI-Beispiele

Unter Berücksichtigung der quantitativen Auswertungen auf Basis der Mittelwertsvergleiche und der qualitativen Aussagen sowie im Abgleich mit den Befunden aus der Literatur, mittels derer die jeweiligen Beispiele aufgrund der Häufigkeit des Auftretens des Krankheitsbildes in der Bevölkerung als relevant eingestuft werden können, wurden die nachfolgenden KI-Beispiele als am geeignetsten für die Akzeptanzstudie identifiziert:

- a) Prävention: Bildbasiertes Erkennen von Darmpolypen
- b) Diagnostik: Sensorbasiertes Erkennen von Herz-Kreislauf-Erkrankungen
- c) Therapie: Sensorbasiertes Erkennen von Über- bzw. Unterzuckerung

Abbildungen 9-11 erläutern den Auswahlprozess.

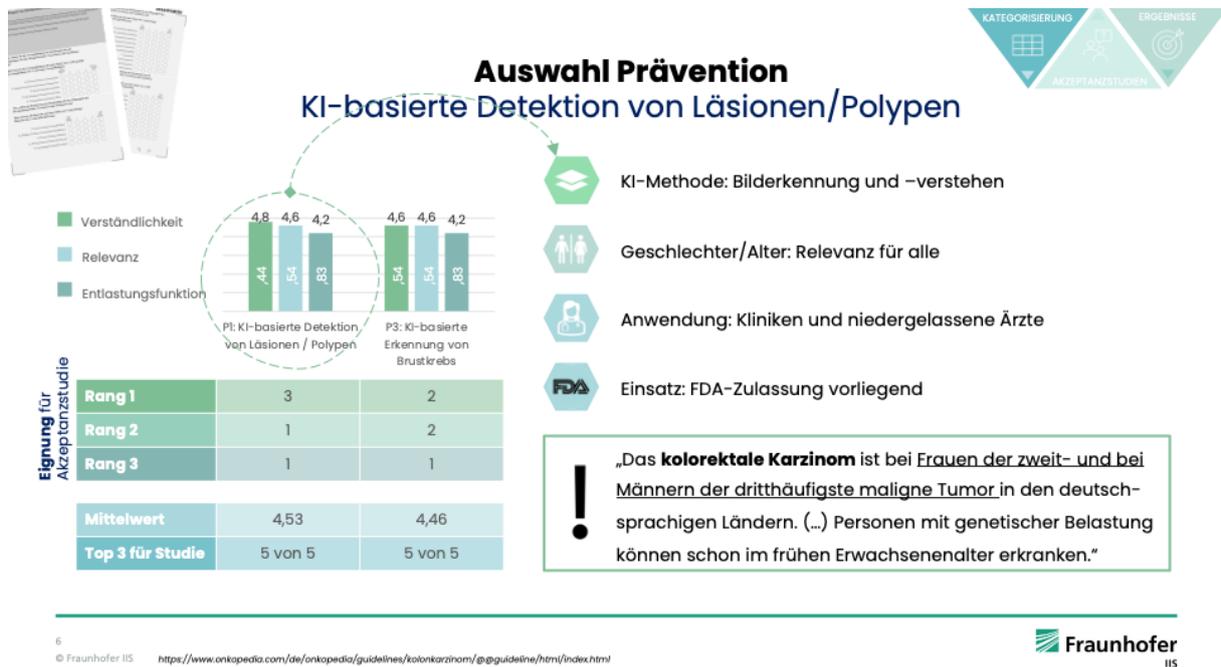


Abbildung 9. Ableitung der Auswahl des KI-Beispiels - Bereich Prävention

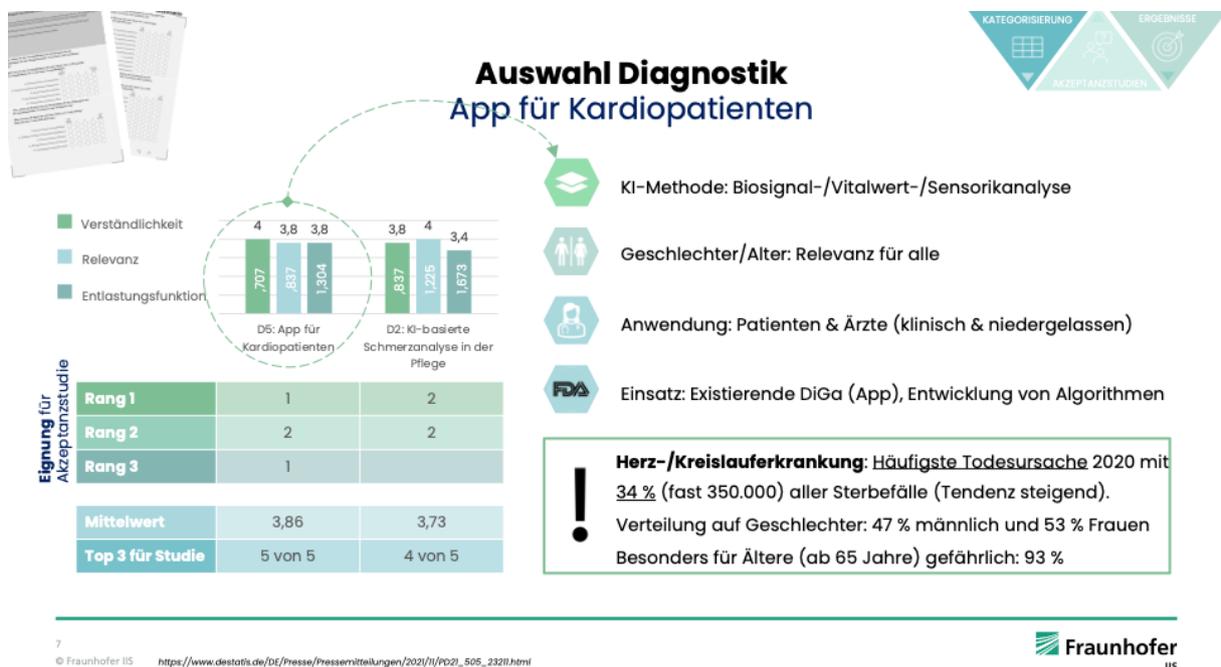


Abbildung 10. Ableitung der Auswahl des KI-Beispiels - Bereich Diagnostik



Auswahl Therapie Frühzeitige Erkennung von Unterzuckerung

- Verständlichkeit
- Relevanz
- Entlastungsfunktion



Eignung für Akzeptanzstudie	T8: App für Kardiopatienten	T4: Frühzeitige Erkennung von Unterzuckerung auf der Intensivstation
Rang 1	1	1
Rang 2	1	2
Rang 3	2	
Mittelwert	3,89	4,00
Top 3 für Studie	5 von 5	3 von 5

- KI-Methoden:** Biosignal-/Vitalwert-/Sensorikanalyse
Mensch-Maschine-Interaktion
- Geschlechter/Alter:** Relevanz für alle
- Anwendung:** Patienten & Ärzte (klinisch & niedergelassen)
- FDA Einsatz:** Bereits in Anwendung (klinischer Bereich)
Wünschenswerte Entwicklung (Privatgebrauch)

8 Mio. Diabetiker Typ-2 in D 2020 (+ Dunkelziffer von 2 Mio.), bis zu **600.000 Neuerkrankungen jährlich** → 2040 etwa 12 Mio.; bis zu **2,6-faches Risiko für frühzeitigen Tod**, 5-6 Jahre geringere Lebenserwartung, **zahlreiche Auswirkungen auf Augen, Nieren, Nerven und Herz**; bei stationären Fällen **hat jeder fünfte Diabetes**

© Fraunhofer IIS https://www.deutsche-diabetes-gesellschaft.de/fileadmin/user_upload/06_Gesundheitspolitik/03_Veroeffentlichungen/05_Gesundheitsbericht/2020/107_Gesundheitsbericht2021.pdf (S.8-12)
<https://www.aerzteblatt.de/nachrichten/123312/Jeder-fuenfte-stationaere-Patient-hat-Diabetes>

Abbildung 11. Ableitung der Auswahl des KI-Beispiels - Bereich Therapie

Literatur

- Atanasov, I., Gauthier, S., & Lopes, C. (2018). Applications of artificial intelligence technologies in healthcare: A systematic literature review. *Journal of Artificial Intelligence Research*, 61, 279-314.
- Ben-Israel, O., Jacobs, D., Casha, S., Lang, S., Ryu, W. H., de Lotbiniere-Bassett, M., & Caddotte, D. W. (2020). The impact of machine learning on patient care: A systematic review. *Artificial Intelligence in Medicine*, 104, 101789.
- Boell, S. K., & Cecez-Kecmanovic, D. (2015). On being 'systematic' in literature reviews in IS. *Journal of Information Technology*, 30(2), 161-173.
- Cooper, H. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society*, 1(1), 104-126.
- Cooper, H., & Hedges, L. V. (2009). Research synthesis as a scientific process. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 3-16). Russell Sage Foundation.
- Fettke, P. (2006). State-of-the-art des State-of-the-Art: Eine Untersuchung der Forschungsmethode „Review“ innerhalb der Wirtschaftsinformatik. *Wirtschaftsinformatik*, 48(4), 257-266.
- Fink, A. (2014). *Conducting research literature reviews: From the Internet to paper* (4th ed.). SAGE Publications.
- Kitchenham, B., & Charters, S. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering*. EBSE Technical Report.
- Levy, Y., & Ellis, T. J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science: International Journal of an Emerging Transdiscipline*, 9, 181-212.
- Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12. Aufl.). Weinheim: Beltz.
- Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Malden, MA: Blackwell Publishing.
- Rothstein, H. R., & Hopewell, S. (2009). Grey literature. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 103-125). Russell Sage Foundation.
- Schryen, G. (2015). Writing qualitative IS literature reviews – Guidelines for synthesis, interpretation, and guidance of research. *Communications of the AIS*, 37, 286-325.
- vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., & Cleven, A. (2009). Reconstructing the giant: On the importance of rigour in documenting the literature search process. In *Proceedings of the 17th European Conference on Information Systems* (pp. 2206-2217).

vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the Association for Information Systems*, 37(1), 205-224.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, 26(2), xiii-xxiii.

Wolff, J., Pauling, J., Keck, A., & Baumbach, J. (2020). The economic impact of artificial intelligence in healthcare: Systematic review. *Health Economics Review*, 10(1), 27.

Zhang, H., Babar, M. A., & Tell, P. (2011). Identifying relevant studies in software engineering. *Information and Software Technology*, 53(6), 625-637.

Wolfswinkel, J. F., Furtmueller, E., & Wilderom, C. P. M. (2013). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, 22(1), 45-55.

Anhang

Anhang 1: Online-Fragebogen zur Stimulusauswahl auf Basis des Kategorisierungsinstrument



Wir vom Fraunhofer-Institut für Integrierte Schaltungen IIS entwickeln derzeit im vom Innovationsausschuss des G-BA geförderten Projekt "KI-BA - Künstliche Intelligenz in der Versorgung - Bedingungen der Akzeptanz von Versicherten und ÄrztInnen" ein mehrdimensionales Kategorisierungssystem für KI in der medizinischen Versorgung.

Bei der Entwicklung des Kategorisierungssystems hast du uns Ende 2021 im Rahmen eines Experteninterviews sehr unterstützt und uns eine Reihe von KI-Beispielen genannt.

Wie in den Interviews schon angekündigt, würden wir uns sehr freuen, wenn du uns nun noch einmal mit deinem Expertenwissen unterstützt:

Wir möchten verstehen, welche KI-Beispiele aus deiner Sicht am besten geeignet sind, um die Akzeptanz von KI mit einer sehr heterogenen Zielgruppe von ca. 1500 Versicherten und 500 ÄrztInnen unterschiedlichen Alters zu erheben - egal, ob es die KI-Beispiele bereits am Markt gibt oder ob sie sich noch in der Entwicklung befinden. Wir möchten verstehen, welche KI-Beispiele du selbst an dir anwenden lassen würdest - egal, ob es die KI-Beispiele bereits am Markt gibt oder ob sie sich noch in der Entwicklung befinden. Wir möchten verstehen, welche KI-Beispiele du als Arzt verwenden würdest - egal, ob es die KI-Beispiele bereits am Markt gibt oder ob sie sich noch in der Entwicklung befinden. Wir möchten verstehen, wie die KI gestaltet sein müsste, damit du sie an dir anwenden lassen bzw. selbst verwenden würdest - egal, ob es die KI bereits am Markt gibt oder ob sie sich noch in der Entwicklung befindet.

Was erwartest dich?

Wir werden dir alle KI-Beispiele zeigen, die in unseren 7 Experteninterviews genannt wurden - gegliedert in die Versorgungssäulen Prävention, Diagnostik und Therapie. Im Anschluss an die Beispiele bitten wir dich jeweils um deine Einschätzung der KI-Beispiele. Die Befragung ist vollständig anonym. Wir werden dich zu keinem Zeitpunkt nach Daten fragen, die Rückschlüsse auf deine Person zulassen. Die Ausfüllzeit beträgt ca. 15 Minuten.

Schon jetzt ganz herzlichen Dank für deine erneute Unterstützung und viel Spaß beim Ausfüllen!

Wir sind gespannt auf deine Einschätzung und Anregungen!



Teil A: KI-Beispiele aus dem Bereich Prävention (1)

Im Folgenden findest du eine Übersicht aller fünf KI-Beispiele aus dem Bereich Prävention (P1-P5), die im Rahmen unserer Experteninterviews zum KI-Kategorisierungssystem genannt und in das Kategoriensystem eingeordnet wurden.

Bitte sieh dir die fünf Beispiele und ihre jeweils erläuternde Kategorisierung an und beantworte dann die Fragen.

Zur Beantwortung der Fragen kannst du dir die Beispiele jederzeit ansehen.

A1. Wie schätzt du die Verständlichkeit der KI-Beispiele für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) ein?

Bitte bewerte die Verständlichkeit auf einer Skala von 1 (sehr geringe Verständlichkeit) bis 5 (sehr hohe Verständlichkeit).

	1 (sehr geringe Verständlichkeit)	2	3	4	5 (sehr hohe Verständlichkeit)
P1: KI-basierte Detektion von Läsionen/Polypen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
P2: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
P3: KI-basierte Erkennung von Brustkrebs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
P4: GKV-Routinedaten für Parkinson-Sturz-Prävention	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
P5: Verletzungsdiagnose paralympischer Athleten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

A2. Wie schätzt du die Relevanz der KI-Beispiele für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) ein?

Bitte bewerte die Relevanz auf einer Skala von 1 (sehr geringe Relevanz) bis 5 (sehr hohe Relevanz).

	1 (sehr geringe Relevanz)	2	3	4	5 (sehr hohe Relevanz)
P1: KI-basierte Detektion von Läsionen/Polypen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
P2: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
P3: KI-basierte Erkennung von Brustkrebs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
P4: GKV-Routinedaten für Parkinson-Sturz-Prävention	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
P5: Verletzungsdiagnose paralympischer Athleten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



A3. Wie schätzt du die Entlastungsfunktion der KI-Beispiele für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) ein?

Bitte bewerte die Entlastungsfunktion auf einer Skala von 1 (sehr geringe Entlastung) bis 5 (sehr hohe Entlastung).

	1 (sehr geringe Entlastung)	2	3	4	5 (sehr hohe Entlastung)
P1: KI-basierte Detektion von Läsionen/Polypen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
P2: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
P3: KI-basierte Erkennung von Brustkrebs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
P4: GKV-Routinedaten für Parkinson-Sturz-Prävention	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
P5: Verletzungsdiagnose paralympischer Athleten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Teil B: KI-Beispiele aus dem Bereich Prävention (2)

Im Folgenden findest du nochmals die Übersicht aller fünf KI-Beispiele aus dem Bereich Prävention (P1-P5), die im Rahmen unserer Experteninterviews zum KI-Kategorisierungssystem genannt und in das Kategoriensystem eingeordnet wurden.

Bitte sieh dir die fünf Beispiele und ihre jeweils erläuternde Kategorisierung an und beantworte dann die Fragen.

Zur Beantwortung der Fragen kannst du dir die Beispiele jederzeit ansehen.

B1. Nachdem du nun alle Beispiele im Bereich "Prävention" eingeschätzt hast: Welche drei Beispiele sind aus deiner Sicht am besten geeignet, um sie gleichzeitig mit einer großen Anzahl von Versicherten und ÄrztInnen unterschiedlichen Alters, Bildungsgrades, Technikaffinität etc. bezüglich KI-Akzeptanz zu testen?

Bitte ziehe deine drei Favoriten auf die linke Seite, beginnend mit dem aus deiner Sicht am besten geeigneten Beispiel.

P1: KI-basierte Detektion von Läsionen/Polypen	<input type="checkbox"/>
P2: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	<input type="checkbox"/>
P3: KI-basierte Erkennung von Brustkrebs	<input type="checkbox"/>
P4: GKV-Routinedaten zur Parkinson-Sturz-Prävention	<input type="checkbox"/>
P5: Verletzungsdiagnose paralympischer Athleten	<input type="checkbox"/>



B2. Du hast nun für den Bereich Prävention die aus deiner Sicht am besten geeigneten KI-Beispiele für die Akzeptanzstudie mit Versicherten und ÄrztInnen ausgewählt. Bitte nenne uns kurz die Gründe für deine Auswahl.

B3. Versetze dich in die Rolle eines Patienten/einer Patientin:

Bezogen auf die von dir ausgewählten drei Beispiele: In welchem Fall würdest du die KI in deinem konkreten Beispiel an dir anwenden lassen?

Wähle pro Beispiel den jeweils am besten zutreffenden Fall aus.

	Würde ich nicht an mir anwenden lassen.	Würde ich es nur annehmen, wenn die KI nur beschriebene Fälle (z.B. eine kleine Einkerbung auf dem Zahn) betrifft.	Würde ich es nur annehmen, wenn die KI kompetent ist und ich mich nicht selbst um die Sache kümmern muss.	Würde ich es mir anwenden lassen, wenn die KI selbst eine Entscheidung trifft (z.B. ob ich einen Zahn ziehen muss).
Mein am besten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mein am zweitbesten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mein am drittbesten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

B4. Als Patient/Patientin:

Würdest du der KI in deinen Beispielen die alleinige Entscheidung überlassen, wenn dir zusätzliche Informationen zur Verfügung stünden?

Ja
Nein

B5. Als Patient/Patientin:

Welche zusätzlichen Informationen würdest du dir wünschen?



B6. Versetze dich in die Rolle eines Arztes/einer Ärztin:

Bezogen auf die von dir ausgewählten drei Beispiele: In welchem Fall würdest du die KI in deinem konkreten Beispiel verwenden?

Wähle pro Beispiel den jeweils am besten zutreffenden Fall aus.

	Würde ich nicht verwenden.	Würde ich verwenden, wenn die KI ein beschriebenes oder fiktives, aber keine Stimmbelegungsfall (Stimm in the loop).	Würde ich verwenden, wenn die KI interpretiert oder analysiert und Vorschläge macht (Stimm on the loop).	Würde ich verwenden, wenn die KI selbst entscheidet oder handelt (Stimm out of the loop).
Mein am besten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mein am zweitbesten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mein am drittbesten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

B7. Als Arzt/Ärztin:

Würdest du der KI in deinen Beispielen die alleinige Entscheidung überlassen, wenn dir zusätzliche Informationen zur Verfügung stünden?

Ja

Nein

B8. Als Arzt/Ärztin:

Welche zusätzlichen Informationen würdest du dir wünschen?



Teil C: KI-Beispiele aus dem Bereich Diagnostik (1)

Im Folgenden findest du eine Übersicht aller acht KI-Beispiele aus dem Bereich Diagnostik (D1-D8), die im Rahmen unserer Experteninterviews zum KI-Kategorisierungssystem genannt und in das Kategoriensystem eingeordnet wurden.

Bitte sieh dir die acht Beispiele und ihre jeweils erläuternde Kategorisierung an und beantworte dann die Fragen.

Zur Beantwortung der Fragen kannst du dir die Beispiele jederzeit ansehen.

C1. Wie schätzt du die Verständlichkeit der KI-Beispiele für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) ein?

Bitte bewerte die Verständlichkeit auf einer Skala von 1 (sehr geringe Verständlichkeit) bis 5 (sehr hohe Verständlichkeit).

	1 (sehr geringe Verständlichkeit)	2	3	4	5 (sehr hohe Verständlichkeit)
D1: Multimodale Zeitreihendaten (Fitnessshirt-Monitoring)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D2: KI-basierte Schmerzanalyse in der Pflege	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D3: Ultraschallbasierte Flüssigkeitsgabe im Notarzttwesen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D4: Blutbildanalyse Hämacam	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D5: App für Kardiopatienten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D6: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D7: Digitale Pathologie (medizinische Bildanalyse)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D8: Leitlinienmodellierung / longitudinale Modellierung	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



C2. Wie schätzt du die Relevanz der KI-Beispiele für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) ein?

Bitte bewerte die Relevanz auf einer Skala von 1 (sehr geringe Relevanz) bis 5 (sehr hohe Relevanz).

	1 (sehr geringe Relevanz)	2	3	4	5 (sehr hohe Relevanz)
D1: Multimodale Zeitreihendaten (Fitnessshirt-Monitoring)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D2: KI-basierte Schmerzanalyse in der Pflege	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D3: Ultraschallbasierte Flüssigkeitsgabe im Notarztwesen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D4: Blutbildanalyse Hämacam	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D5: App für Kardiopatienten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D6: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D7: Digitale Pathologie (medizinische Bildanalyse)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D8: Leitlinienmodellierung / longitudinale Modellierung	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

C3. Wie schätzt du die Entlastungsfunktion der KI-Beispiele für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) ein?

Bitte bewerte die Entlastungsfunktion auf einer Skala von 1 (sehr geringe Entlastung) bis 5 (sehr hohe Entlastung).

	1 (sehr geringe Entlastung)	2	3	4	5 (sehr hohe Entlastung)
D1: Multimodale Zeitreihendaten (Fitnessshirt-Monitoring)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D2: KI-basierte Schmerzanalyse in der Pflege	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D3: Ultraschallbasierte Flüssigkeitsgabe im Notarztwesen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D4: Blutbildanalyse Hämacam	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D5: App für Kardiopatienten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D6: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D7: Digitale Pathologie (medizinische Bildanalyse)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D8: Leitlinienmodellierung / longitudinale Modellierung	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



Teil D: KI-Beispiele aus dem Bereich Diagnostik (2)

Im Folgenden findest du nochmals Übersicht aller acht KI-Beispiele aus dem Bereich Diagnostik (D1-D8), die im Rahmen unserer Experteninterviews zum KI-Kategorisierungssystem genannt und in das Kategoriensystem eingeordnet wurden.

Bitte sieh dir die acht Beispiele und ihre jeweils erläuternde Kategorisierung an und beantworte dann die Fragen.

Zur Beantwortung der Fragen kannst du dir die Beispiele jederzeit ansehen.

D1. Nachdem du nun alle Beispiele im Bereich "Diagnostik" eingeschätzt hast: Welche drei Beispiele sind aus deiner Sicht am besten geeignet, um sie gleichzeitig mit einer großen Anzahl von Versicherten und ÄrztInnen unterschiedlichen Alters, Bildungsgrades, Technikaffinität etc. bezüglich KI-Akzeptanz zu testen?

Bitte ziehe deine drei Favoriten auf die linke Seite, beginnend mit aus deiner Sicht am besten geeigneten Beispiel.

D1: Multimodale Zeitreihendaten (Fitnessshirt-Monitoring)

D2: KI-basierte Schmerzanalyse in der Pflege

D3: Ultraschallbasierte Flüssigkeitsgabe im Notarzteswesen

D4: Blutbildanalyse Hämacam

D5: App für Kardiopatienten

D6: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten

D7: Digitale Pathologie (medizinische Bildanalyse)

D8: Leitlinienmodellierung/longitudinale Modellierung

D2. Du hast nun für den Bereich Diagnostik die aus deiner Sicht am besten geeigneten KI-Beispiele für die Akzeptanzstudie mit Versicherten und ÄrztInnen ausgewählt. Bitte nenne uns kurz die Gründe für deine Auswahl.



D3. Versetze dich in die Rolle eines Patienten/einer Patientin:

Bezogen auf die von dir ausgewählten drei Beispiele: In welchem Fall würdest du die KI in deinem konkreten Beispiel an dir anwenden lassen?

Wähle pro Beispiel den jeweils am besten zutreffenden Fall aus.

	Würde ich nicht an mir anwenden lassen.	Würde ich mir anwenden lassen, wenn die KI nur beobachtet oder FMR, aber keine Entscheidung trifft. (Human in the loop).	Würde ich mir anwenden lassen, wenn die KI interpretiert oder analysiert und Vorschläge macht. (Human in the loop).	Würde ich mir anwenden lassen, wenn die KI selbst entscheidet oder handelt (Human out of the loop).
Mein am besten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mein am zweitbesten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mein am drittbesten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

D4. Als Patient/Patientin:

Würdest du der KI in deinen Beispielen die alleinige Entscheidung überlassen, wenn dir zusätzliche Informationen zur Verfügung stünden?

Ja
Nein

D5. Als Patient/Patientin:

Welche zusätzlichen Informationen würdest du dir wünschen?

D6. Versetze dich in die Rolle eines Arztes/einer Ärztin:

Bezogen auf die von dir ausgewählten drei Beispiele: In welchem Fall würdest du die KI in deinem konkreten Beispiel verwenden?

Wähle pro Beispiel den jeweils am besten zutreffenden Fall aus.

	Würde ich nicht verwenden.	Würde ich verwenden, wenn die KI nur beobachtet oder FMR, aber keine Entscheidung trifft. (Human in the loop).	Würde ich verwenden, wenn die KI interpretiert oder analysiert und Vorschläge macht. (Human in the loop).	Würde ich verwenden, wenn die KI selbst entscheidet oder handelt (Human out of the loop).
Mein am besten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mein am zweitbesten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mein am drittbesten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



D7. Als Arzt/Ärztin:

Würdest du der KI in deinen Beispielen die alleinige Entscheidung überlassen, wenn dir zusätzliche Informationen zur Verfügung stünden?

Ja

Nein

D8. Als Arzt/Ärztin:

Welche zusätzlichen Informationen würdest du dir wünschen?



Teil E: KI-Beispiele aus dem Bereich Therapie (1)

Im Folgenden findest du eine Übersicht aller zehn KI-Beispiele aus dem Bereich Therapie (T1-T10), die im Rahmen unserer Experteninterviews zum KI-Kategorisierungssystem genannt und in das Kategoriensystem eingeordnet wurden.

Bitte sieh dir die zehn Beispiele und ihre jeweils erläuternde Kategorisierung an und beantworte dann die Fragen.

Zur Beantwortung der Fragen kannst du dir die Beispiele jederzeit ansehen.

E1. Wie schätzt du die Verständlichkeit der KI-Beispiele für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) ein?

Bitte bewerte die Verständlichkeit auf einer Skala von 1 (sehr geringe Verständlichkeit) bis 5 (sehr hohe Verständlichkeit).

	1 (sehr geringe Verständlichkeit)	2	3	4	5 (sehr hohe Verständlichkeit)
T1: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T2: Genetisch spezifizierte Antibiotikagabe im Krankenhaus	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T3: Ultraschallbasierte Flüssigkeitsgabe im Notarztesen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T4: Frühzeitige Erkennung von Unterzuckerung auf der Intensivstation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T5: Frühzeitige Erkennung von Unterzuckerung für den Privatgebrauch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T6: Longitudinale Datenauswertung/Krankheitsprädiaktion	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T7: Autismustherapie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T8: App für Kardiopatienten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T9: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T10: Leitlinienmodellierung/Longitudinale Modellierung	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



E2. Wie schätzt du die Relevanz der KI-Beispiele für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) ein?

Bitte bewerte die Relevanz auf einer Skala von 1 (sehr geringe Relevanz) bis 5 (sehr hohe Relevanz).

	1 (sehr geringe Relevanz)	2	3	4	5 (sehr hohe Relevanz)
T1: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T2: Genetisch spezifizierte Antibiotikagabe im Krankenhaus	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T3: Ultraschallbasierte Flüssigkeitsgabe im Notarztwesen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T4: Frühzeitige Erkennung von Unterzuckerung auf der Intensivstation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T5: Frühzeitige Erkennung von Unterzuckerung für den Privatgebrauch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T6: Longitudinale Datenauswertung/Krankheitsprädiktion	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T7: Autismustherapie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T8: App für Kardiopatienten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T9: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T10: Leitlinienmodellierung/Longitudinale Modellierung	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

E3. Wie schätzt du die Entlastungsfunktion der KI-Beispiele für die Zielgruppen der KI-Akzeptanzstudie (Versicherte und ÄrztInnen) ein?

Bitte bewerte die Entlastungsfunktion auf einer Skala von 1 (sehr geringe Entlastung) bis 5 (sehr hohe Entlastung).

	1 (sehr geringe Entlastung)	2	3	4	5 (sehr hohe Entlastung)
T1: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T2: Genetisch spezifizierte Antibiotikagabe im Krankenhaus	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T3: Ultraschallbasierte Flüssigkeitsgabe im Notarztwesen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T4: Frühzeitige Erkennung von Unterzuckerung auf der Intensivstation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T5: Frühzeitige Erkennung von Unterzuckerung für den Privatgebrauch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T6: Longitudinale Datenauswertung/Krankheitsprädiktion	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T7: Autismustherapie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T8: App für Kardiopatienten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T9: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T10: Leitlinienmodellierung/Longitudinale Modellierung	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



Teil F: KI-Beispiele aus dem Bereich Therapie (2)

Im Folgenden findest du nochmals die Übersicht aller zehn KI-Beispiele aus dem Bereich Therapie (T1-T10), die im Rahmen unserer Experteninterviews zum KI-Kategorisierungssystem genannt und in das Kategoriensystem eingeordnet wurden.

Bitte sieh dir die zehn Beispiele und ihre jeweils erläuternde Kategorisierung an und beantworte dann die Fragen.

Zur Beantwortung der Fragen kannst du dir die Beispiele jederzeit ansehen.

F1. Nachdem du nun alle Beispiele im Bereich "Therapie" eingeschätzt hast: Welche drei Beispiele sind aus deiner Sicht am besten geeignet, um sie gleichzeitig mit einer großen Anzahl von Versicherten und ÄrztInnen unterschiedlichen Alters, Bildungsgrades, Technikaffinität etc. bezüglich KI-Akzeptanz zu testen?

Bitte ziehe deine drei Favoriten auf die linke Seite, beginnend mit aus deiner Sicht am besten geeigneten Beispiel.

T1: AR-Brillen in der Pflege zur Unterstützung des Morgenprozesses	<input type="checkbox"/>
T2: Genetisch spezifizierte Antibiotikagabe im Krankenhaus	<input type="checkbox"/>
T3: Ultraschallbasierte Flüssigkeitsgabe im Notarztesen	<input type="checkbox"/>
T4: Frühzeitige Erkennung von Unterzuckerung auf der Intensivstation	<input type="checkbox"/>
T5: Frühzeitige Erkennung von Unterzuckerung für den Privatgebrauch	<input type="checkbox"/>
T6: Longitudinale Datenauswertung/Krankheitsprädiktion	<input type="checkbox"/>
T7: Autismustherapie	<input type="checkbox"/>
T8: App für Kardiopatienten	<input type="checkbox"/>
T9: Gangtest-Objektivierung durch Schuhsensoren für Parkinsonpatienten	<input type="checkbox"/>
T10: Leitlinienmodellierung/longitudinale Modellierung	<input type="checkbox"/>



F2. Du hast nun für den Bereich Therapie die aus deiner Sicht am besten geeigneten KI-Beispiele für die Akzeptanzstudie mit Versicherten und ÄrztInnen ausgewählt. Bitte nenne uns kurz die Gründe für deine Auswahl.

F3. Versetze dich in die Rolle eines Patienten/einer Patientin:

Bezogen auf die von dir ausgewählten drei Beispiele: In welchem Fall würdest du die KI in deinem konkreten Beispiel an dir anwenden lassen?

Wähle pro Beispiel den jeweils am besten zutreffenden Fall aus.

	Würde ich nicht an mir anwenden lassen.	Würde ich es mir anwenden lassen, wenn die KI zur Beobachtung oder Fälli, aber keine Entscheidung trifft. (Observe in the loop).	Würde ich es mir anwenden lassen, wenn die KI interpretiert oder analysiert und Vorschläge macht. (Assess in the loop).	Würde ich es mir anwenden lassen, wenn die KI selbst entscheidet oder handelt. (Decide and act of the loop).
Mein am besten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mein am zweitbesten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mein am drittbesten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

F4. Als Patient/Patientin:

Würdest du der KI in deinen Beispielen die alleinige Entscheidung überlassen, wenn dir zusätzliche Informationen zur Verfügung stünden?

Ja
Nein

F5. Als Patient/Patientin:

Welche zusätzlichen Informationen würdest du dir wünschen?



F6. Versetze dich in die Rolle eines Arztes/einer Ärztin:

Bezogen auf die von dir ausgewählten drei Beispiele: In welchem Fall würdest du die KI in deinem konkreten Beispiel verwenden?

Wähle pro Beispiel den jeweils am besten zutreffenden Fall aus.

	Würde ich nicht verwenden.	Würde ich verwenden, wenn die KI nur Informationen nicht falsch, aber keine Entscheidung trifft? (Dokument in der Regel)	Würde ich verwenden, wenn die KI Informationen oder mehrheitlich Vorschläge macht (Dokument vor der Regel)	Würde ich verwenden, wenn die KI selbst entscheidet oder bestimmte Daten vor of der Regel)
Mein am besten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mein am zweitbesten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mein am drittbesten für die Akzeptanzstudie geeignetes Beispiel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

F7. Als Arzt/Ärztin:

Würdest du der KI in deinen Beispielen die alleinige Entscheidung überlassen, wenn dir zusätzliche Informationen zur Verfügung stünden?

Ja
Nein

F8. Als Arzt/Ärztin:

Welche zusätzlichen Informationen würdest du dir wünschen?

Geschafft!

Herzlichen Dank für deine Teilnahme!

Anhang 2: AP 1 – Konzeptmatrizen und verwendete Literatur

Input

Kategorie Input	Ausprägung Zeitreihendaten	Ausprägung Bildgebungsdaten	Ausprägung Audiodaten	Ausprägung Daten aus Untersuchung, Labor, Demografie	Ausprägung Sonstige klinische Daten (Handschriftliche Notizen, Begegnungen)	Ausprägung "omics"-Daten	Ausprägung Weitere
L020		x	x		x		+
L005	x	x		x	x	x	
L032		x			x		
L036	x	x		x			+
L015	x	x		x		x	
L004	x			x		x	

ID	Link	Titel	Seite	Kapitel	Ausprägung 1	Ausprägung 2	Ausprägung 3	Ausprägung 4	Ausprägung 5	Ausprägung 6	Ausprägung 7	Ausprägung 8	Ausprägung 9	Ausprägung 10	
L020	https://www.nature.com/articles/nature14234	Künstliche Intelligenz – eine Einführung in Technologiefelder & Forschungsbereiche		8	4.2 Datenquellen	Unstrukturiert: Bilder	Unstrukturiert: Video	Unstrukturiert: Text	Unstrukturiert: Audio (Environmental, Speech)	Unstrukturiert: IoT und komplexe System-Sensordaten	Unstrukturiert: Netzwerkdaten	Unstrukturiert: Log-files von Software-Applikationen	Unstrukturiert: QoS Parameter von Netzwerkprotokollen (network traffic)	Strukturiert: Datenbanken	Strukturiert: CSV files
L005	doi:10.1016/j.artmed.2019.101785	The impact of machine learning on patient care: a systematic review. Artificial Intelligence in Medicine, 101785.		2	2.1.3 Data classification	Unstrukturierte Zeitreihendaten als klinische Daten	Unstrukturierte Bilddgebungsdaten als radiologische Rohbilder sowie medizinisch relevante Digitalfotos	Unstrukturierte sonstige klinische Daten wie handschriftliche Notizen oder diktierte klinische Begegnungen	Strukturierte klinische (körperliche Untersuchungen etc.), demografische und Labordaten als klinische Daten	Strukturierte genomische Daten als digitalisierte genomische Sequenzen					
L032	https://link.springer.com	Blockchain and artificial intelligence technology in e-Health		52811	Introduction	Unstrukturierte Daten (Patient, Radiologie, Bilder)									
L036	https://eeexplore.ieee.org	A Review on Deep Learning Applications in Prognostics and Health Management		162416	Introduction	Vibration (inkl. Akustik) - Daten	Bilddgebungsdaten	Strukturierte Daten	Zeitreihendaten						
L015	https://cdt.amegroups.com	Artificial intelligence in personalized cardiovascular medicine and cardiovascular imaging		911-912	Why we need artificial intelligence	"omics"-Daten	Bilddgebungsdaten	Daten von Sensoren & Wearables	Daten der ePA (Electronic health records)						
L004	https://link.springer.com	Big Data und künstliche Intelligenz in der Medizin (in: Telemedizin - Grundlagen und praktische Anwendung in stationären und ambulanten Krankheitszuständen)		425-429	Überblick über Big-Data-Analysen	"omics"-Daten	Strukturierte Daten (stationär wie Geschlecht und Körper + dynamisch Temperatur, Werte)	Unstrukturierte Daten aus elektronischen Krankenakten	Strukturierte dynamische Daten von Sensoren & Wearables						

Nicht IT-bezogen:

https://www.plattform-itn.de	Sichere KI-Systeme für die Medizin - Datenmanagement und IT-Sicherheit in der	11	„Mit KI gegen Krebs“, Verbesserte Diagnose und	anatomische Informationen	physiologische Informationen	mechanische Informationen	pathologische Informationen	chirurgische Informationen	metabolische Informationen	genetische Informationen	radiologische Informationen	soziologische Werte	psychologische Werte
https://www.msdfman.com/de	MSD Manual (Internet)		Untersuchungsmethoden	Körperflüssigkeiten	Bilddgebende Verfahren	Endoskopie	Körperfunktionen	Biopsie	Genetische Untersuchung				
https://www.thieme.de	i care Pflege - Thieme Verlag		Diagnostische Verfahren	Verfahren ohne apparative Hilfsmittel (Anamnese, Inspektion, Palpation)	Verfahren mit apparativen Hilfsmitteln (Blutuntersuchung, Röntgen, CT, MRT)	Nicht-invasive Diagnostik	Invasive Diagnostik						
https://books.google.de/books?id=8U...	Checkliste innere Medizin: 266 Tabellen - Johannes-Martin Hahn - Thieme Verlag	VII	Inhaltsverzeichnis	Laboruntersuchungen	Nicht-invasive Diagnostik	Invasive Diagnostik und Therapie	Injektions- und Funktionstechniken						

Versorgungsäulen

Kategorie Versorgungssäule	Ausprägung Prävention	Ausprägung Diagnostik	Ausprägung Therapie	Ausprägung Rehabilitation	Ausprägung Pflege	Ausprägung Weitere
L004	X	X	X	X	X	
L008	X	X	X	X	X	+
L009	X	X	X	X	X	+
L017	X			X	X	
L011	X	X	X	X		+
L030	X	X	X	X	X	
L033	X			X	X	+
L028	X			X	X	+
L006	X	X	X	X		+
L016	X		X	X	X	+
L027	X			X		+
L007	X	X	X	X		+
L034	X	X	X			+
L019		X	X		X	+

Nr.	Link	Titel	Seite	Kapitel	Ausprägung1	Ausprägung2	Ausprägung3	Ausprägung4	Ausprägung5	Ausprägung6	Ausprägung7	Ausprägung8
L004	https://www.gesundheitswissenschaften.de/DT/B/Technologien.de/DT/B/Praktion/DE/Download/98 https://www.gesundheitswissenschaften.de/DT/B/Technologien.de/DT/B/Praktion/DE/Download/98	Anwendung Künstlicher Intelligenz in der Medizin	4	Management Summary	Prävention	Diagnose	Therapie	Rehabilitation	Pflege			
L008	https://www.gesundheitswissenschaften.de/DT/B/Technologien.de/DT/B/Praktion/DE/Download/98	Digitalisierung in der Medizintechnik	4	Von digitaler Diagnostik zur digitalen Therapie	Prävention	Diagnose	Therapie	Nachsorge	Rehabilitation	Pflege		
L009	https://www.gesundheitswissenschaften.de/DT/B/Technologien.de/DT/B/Praktion/DE/Download/98	Patientenversorgung verbessern – Innovationskraft stärken Fachprogramm	5	Vom nationalen Strategieprozess zum Fachprogramm Medizintechnik	Prävention	Diagnose	Therapie	Nachsorge	Rehabilitation	Pflege		
L017	https://www.gesundheitswissenschaften.de/DT/B/Technologien.de/DT/B/Praktion/DE/Download/98	Versorgungsketten – Der Patient im Mittelpunkt	5	Versorgungseinrichtungen und –ebenen	Prävention	Ambulante Versorgung	Stationäre Versorgung	Rehabilitation	Pflege			
L011	https://www.gesundheitswissenschaften.de/DT/B/Technologien.de/DT/B/Praktion/DE/Download/98	Public Health Kompakt	117	"Gesundheitssysteme"	Gesundheitsförderung und Prävention	Diagnostik und Behandlung von Krankheiten und Unfällen	Rehabilitation von kranken und verunfallten Menschen	gesundheitsbezogene Forschung (i.w.S.)	pharmazeutische Industrie (i.w.S.)	Medizintechnik (i.w.S.)		
L030	https://www.gesundheitswissenschaften.de/DT/B/Technologien.de/DT/B/Praktion/DE/Download/98	Regionen und Netzwerke: Kooperationsmodelle zur	324	"Gesundheit als Standortfaktor"	Prävention	Diagnose	Therapie	Rehabilitation	Pflege			
L033	https://www.gesundheitswissenschaften.de/DT/B/Technologien.de/DT/B/Praktion/DE/Download/98	Unsere Gesundheitssysteme (Wildner, Thieme Verlag)	58	Sektoren des Gesundheitswesens und der professionelle „Mix“	Prävention	Kuration		Rehabilitation	Pflege			
L028	https://www.gesundheitswissenschaften.de/DT/B/Technologien.de/DT/B/Praktion/DE/Download/98	Gesundheit in Deutschland 2015	421-424	8.3 INANSPRUCHNAHME VON LEISTUNGEN DES GESUNDHEITSSYSTEMS	Präventive Leistungen	Medikamentengebrauch	Ambulante Behandlung	Stationäre Behandlung	Rehabilitation	Pflege		
L006	https://www.gesundheitswissenschaften.de/DT/B/Technologien.de/DT/B/Praktion/DE/Download/98	(Muster-)Weiterbildungsordnung 2018	191	Facharzt/Fachärztin für Kinder- und Jugendmedizin	Prävention	Diagnostik	Therapie	Rehabilitation	Nachsorge			
L016	https://www.gesundheitswissenschaften.de/DT/B/Technologien.de/DT/B/Praktion/DE/Download/98	Referenzwerk Prävention und Gesundheitsförderung	29	1.2 Krankheitsprävention und Gesundheitsförderung	Gesundheitsförderung	Prävention	Kuration / Therapie	Rehabilitation	Pflege			
L027	https://www.gesundheitswissenschaften.de/DT/B/Technologien.de/DT/B/Praktion/DE/Download/98	Gesundheitliche Ungleichheit	159	Soziale Ungleichheit und gesundheitsbezogene Ver	Prävention	Kuration	Rehabilitation	Ambulant	Stationär			
L007	https://www.gesundheitswissenschaften.de/DT/B/Technologien.de/DT/B/Praktion/DE/Download/98	Daten helfen heilen	26	Glossar	Ambulante Versorgung	Krankenhaus-Sektor	Rehabilitations-einrichtungen					
L034	https://www.gesundheitswissenschaften.de/DT/B/Technologien.de/DT/B/Praktion/DE/Download/98	Developing an Artificial Intelligence-Enabled Health Care Practice: Rewiring Health Care Professions for Better Care	58	Artificial Intelligence and Health Care	Prävention und öffentliche Gesundheitsfürsorge	Diagnose	Behandlung	Prognose	vorausschauende Gesundheitsdienste			
L019	https://www.gesundheitswissenschaften.de/DT/B/Technologien.de/DT/B/Praktion/DE/Download/98	Factors affecting application of artificial intelligence in healthcare: a scoping review	1	Introduction	Diagnostik	Behandlung	Management	Entscheidungsfindung	Öffentliche Gesundheit	Epidemiologie	Patientenversorgung	Altenpflege

Fachgebiete

Kategorie	Ausprägung Innere Medizin	Ausprägung Neurologie	Ausprägung Intensivmedizin	Ausprägung Chirurgie	Ausprägung Radiologie	Ausprägung Psychiatrie	Ausprägung Urologie	Ausprägung Dermatologie	Ausprägung Pathologie	Ausprägung Pflege	Ausprägung Pharmakologie	Ausprägung Öff. Gesundheit	Ausprägung Weitere
L005	x	x	x	x	x	x	x	x	x	x		x	+
L010	x				x			x	x				
L018	x	x		x	x			x			x		+
L002	x	x	x					x					+
L023	x	x		x	x	x							+
L014	x			x	x				x		x		

Innere Medizin aufgeteilt:

Kategorie	Ausprägung Onkologie	Ausprägung Neurologie	Ausprägung Kardiologie	Ausprägung Intensivmedizin	Ausprägung Chirurgie	Ausprägung Radiologie	Ausprägung Psychiatrie	Ausprägung Urologie	Ausprägung Dermatologie	Ausprägung Endokrinologie	Ausprägung Gastroenterologie	Ausprägung Pathologie	Ausprägung Immunologie	Ausprägung Pflege	Ausprägung Pharmakologie	Ausprägung Öff. Gesundheit	Ausprägung Weitere
Fachgebiete	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	+
	x		x			x			x	x		x					
	x	x	x		x	x			x		x		x		x		+
	x	x		x					x		x		x				+
	x	x	x		x	x	x			x	x		x				+

Nutzer

Kategorie Nutzer	Ausprägung Patienten	Ausprägung Ärzte	Ausprägung Medizinisch/-technisches Fachpersonal	Ausprägung Pflegerisches Fachpersonal	Ausprägung Krankenkasse	Ausprägung Forscher	Ausprägung Angehörige von Patienten	Ausprägung Weitere
L003	X	X	X	X	X	X	X	+
L004	X	X	X	X	X	X	X	+
L012	X	X	X	X	X		X	
L035	X	X	X	X	X	X		
L023		X	X	X	X	X		
L001	X	X	X	X	X			
L019		X	X	X				+

Technologie

Kategorie Technologie	Ausprägung Bildererkennung und verstehen	Ausprägung Datenmanagement und -analyse	Ausprägung Mensch Maschine- Interaktion und Assistenzsysteme	Ausprägung Robotik und autonome Systeme	Ausprägung Sensorik und Kommunikation	Ausprägung Sprach- und Textverstehen	Ausprägung Virtuelle und augmented Realität	Ausprägung Weitere
L022	x	x	x	x	x	x	x	
L021		x	x	x	x			+
L031	x	x		x		x		+
L010			x	x	x	x		
L010	x	x	x		x	x		+
L010			x		x	x	x	+
L030	x	x		x		x		+
L034		x		x		x		+
L029	x	x		x				+

Output

Kategorie Output	Ausprägung Präzisere Diagnose	Ausprägung Personalisierte Behandlung	Ausprägung Verbesserte (Behandlungs-) Abläufe	Ausprägung Verbesserte Beratung	Ausprägung Umfassendere Betreuung	Ausprägung Verbesserter (Versorgungs-) Forschung & Entwicklung	Ausprägung Weitere
L025	x	x			x		+
L004	x	x	x	x	x	x	+
L010			x				
L024	x	x	x	x	x	x	

Beispiele Bildverarbeitende Medizin Personalisierte Medizin (Werte bis Genom) Prozessplanung und -unterstützung, Robotik Vorschlag verschiedener Behandlungsmethoden Nutzerzentrierte DiGas & Apps (Selfcare) Arzneimittelentwicklung, Leitlinien, Auswertungen

Nr.	Link	Titel	Seite	Kapitel	Ausprägung1	Ausprägung2	Ausprägung3	Ausprägung4	Ausprägung5	Ausprägung6	Ausprägung7	Ausprägung8	Ausprägung9	Ausprägung10
L025	https://www.plattform-herz.de/	Sichere KI-Systeme für die Medizin - Datenmanagement	5	Algorithmen	Daten: Effizientere Auswertung (auch von Vitalparametern), Entdeckung von	Diagnostik: Höhere Qualität und Zuverlässigkeit (bessere Entscheidungsmöglichkeiten)	Therapie: Bessere Behandlungsergebnisse durch individuellere Therapien							
L004	https://www.digitale-technologien.de/DIT/Be	Anwendung Künstlicher Intelligenz in der Medizin	10	4 Potenziale und Anwendungen im	neuartige personalisierte Therapien	präzisere Diagnostik und dadurch besserer	effektivere Prävention	Kostensenkung	Erleichterung der Patienten durch weniger Nebenwirkungen und Vermeidung unnötiger	Steigerung der Effizienz	Steigerung der Effektivität	Verbesserung der Gesundheitsversorgung	Personalentlastung bzw. -unterstützung	Patientenstärkung
L010	https://journals.sagepub.com/healthcare-an-essential-guide	Artificial intelligence in healthcare: An essential guide	17	Establishing performance (Effizienz, Sicherheit)	Verbesserte klinische Effektivität (Qualität, Effizienz, Sicherheit)	Verbesserte Patientenerfahrung und -ergebnisse	Erweiterter Zugang und erweiterte Dienstleistungen für Patienten	Geringere Kosten und höhere Einnahmen		Verbesserte Zufriedenheit des Personals mit dem Arbeitsumfeld	Optimierte Betriebsabläufe			
L024	https://www.expona.com/ai-a-short-guide-for-medical-professionals-in-the-era-of	A short guide for medical professionals in the era of	126	EXAMPLES FOR HOW	KI-gesteuerte Beratung (persönlich und online)	Gesundheitliche Betreuung und Medikamentenmanagement	KI-gesteuerte Diagnose	Auswertung von Krankenakten	Präzisionsmedizin	Behandlungspläne	Medikamentenherstellung	Trägerung		

ID	Link	Titel	Seite	Kapitel	Ausprägung1	Ausprägung2	Ausprägung3	Ausprägung4	Ausprägung5	Ausprägung6	Ausprägung7	Ausprägung8	Ausprägung9	Ausprägung10
	https://www.valueinhealthjournal.com/article/S1098-3015(18)32866-3	Value in Health 21 (2018)		Applications of Artificial Intelligence in Healthcare: A Systematic Literature Review	Sensitivität	Spezifität	Genauigkeit							
	10.4103/jmpc.jmpc_44	Overview of artificial intelligence in medicine	2330	Discussion	Effizienz, Genauigkeit, Präzision	Zielerparnis (geringere Arbeitsbelastung, mehr Zeit für Patienten, mehr Zeit für kritische Fälle)	Geldersparnis	Verbesserte Überwachung						

Texte Stimulusvideos – AP 2

Nachfolgend sind die Texte des KI-Einführungsvideos, das gemäß Studiendesign zur Erhebung der KI-Akzeptanz verwendet wurde, sowie der einzelnen KI-Stimuli aufgelistet.

1. Text KI-Einführungsvideo

Künstliche Intelligenz (KI) ist der Überbegriff für Anwendungen, bei denen Maschinen basierend auf Algorithmen menschliche Fähigkeiten wie logisches Denken, Lernen oder Planen ausführen. Sie umfasst dabei verschiedene Anwendungssysteme, von einfachen Algorithmen bis hin zu selbstlernenden Systemen. Künstliche Intelligenz ist heute in vielen Bereichen des täglichen Lebens zu finden: Bei Mobiltelefonen mit Gesichtserkennungstechnologie, digitalen Sprachassistenten oder der Google-Suche.

Auch im Gesundheitswesen hat die künstliche Intelligenz bereits Einzug gehalten: Medizinische Bildanalysen oder roboterassistierte Chirurgie unterstützen Ärzte - von der Behandlung über die Diagnose bis hin zur Dokumentation.

Die zahlreichen Beispiele zeigen es: Technologien, die auf künstlicher Intelligenz beruhen, können viele Formen annehmen: Von Bilderkennung über Sprach- und Textverstehen bis hin zu Biosignalanalysen, Mensch-Maschine-Interaktionen oder Big Data-Analysen. Bei allen in Deutschland am Markt verfügbaren Systemen, die auf künstlicher Intelligenz beruhen, gilt: Datensicherheit und die Einhaltung von Datenschutzrichtlinien sind zentrale Kriterien für die Zulassung und die Nutzung.

2. Übersicht Texte pro KI-Beispiel

2.1 Prävention: Erkennen von Darmpolypen

	AI completeness low	AI completeness high
<p>Explainability low</p>	<p>Zur Darmkrebs-Vorsorge wird unter anderem eine Darmspiegelung des Dickdarms durchgeführt. Hierzu führt der Arzt ein Koloskop, einen biegsamen Schlauch mit einer eingebauten Kamera, in den Darm des Patienten ein. Der Arzt kann so über einen Monitor sehen, ob die Darmschleimhaut gesund aussieht. Er vergleicht hierzu die Monitoransicht mit Normwerten. Er kann so auch tumorverdächtige Polypen erkennen und diese entfernen oder Gewebeprobe entnehmen.</p> <p>Neben dieser Untersuchungsart gibt es Koloskopie-Systeme, die auf künstliche Intelligenz (KI) zurückgreifen: Diese Systeme können selbstständig Darmpolypen über eine Bilderkennung erkennen.</p> <p>Das Koloskopie-System zeigt dem Arzt bei der Untersuchung auf dem Monitor automatisch, wo sich Polypen befinden. Der Arzt kann die weitere Behandlung gemäß des Vorschlags des Systems durchführen. Der Arzt kann auch eine eigene, vom Systemvorschlag abweichende Behandlung durchführen.</p>	<p>Zur Darmkrebs-Vorsorge wird unter anderem eine Darmspiegelung des Dickdarms durchgeführt. Hierzu führt der Arzt ein Koloskop, einen biegsamen Schlauch mit einer eingebauten Kamera, in den Darm des Patienten ein. Der Arzt kann so über einen Monitor sehen, ob die Darmschleimhaut gesund aussieht. Er vergleicht hierzu die Monitoransicht mit Normwerten. Er kann so auch tumorverdächtige Polypen erkennen und diese entfernen oder Gewebeprobe entnehmen.</p> <p>Neben dieser Untersuchungsart gibt es Koloskopie-Systeme, die auf künstliche Intelligenz (KI) zurückgreifen: Diese Systeme können selbstständig Darmpolypen über eine Bilderkennung erkennen.</p> <p>Das Koloskopie-System zeigt dem Arzt bei der Untersuchung auf dem Monitor automatisch, wo sich Polypen befinden und gibt dem Arzt verbindlich die weitere Behandlung vor.</p>
<p>Explainability high</p>	<p>Zur Darmkrebs-Vorsorge wird unter anderem eine Darmspiegelung des Dickdarms durchgeführt. Hierzu führt der Arzt ein Koloskop, einen biegsamen Schlauch mit einer eingebauten Kamera, in den Darm des Patienten ein. Der Arzt kann so über einen Monitor sehen, ob die Darmschleimhaut gesund aussieht. Er vergleicht hierzu die Monitoransicht mit Normwerten. Er kann so auch tumorverdächtige Polypen erkennen und diese entfernen oder Gewebeprobe entnehmen.</p> <p>Neben dieser Untersuchungsart gibt es Koloskopie-Systeme, die auf künstliche Intelligenz (KI) zurückgreifen: Diese Systeme können selbstständig Darmpolypen über eine Bilderkennung erkennen. Die Untersuchung findet dabei weiterhin durch den Arzt statt. Verwendet wird weiterhin ein Koloskop mit einer eingebauten Kamera. Die künstliche Intelligenz greift im Hintergrund auf eine große klinische Datenbank zurück: Sie vergleicht die Daten des Patienten mit den Daten anderer Patienten. Durch diesen Datenabgleich erkennt das System auffällige Gewebeareale und markiert sie farblich.</p> <p>Das Koloskopie-System zeigt dem Arzt bei der Untersuchung auf dem Monitor automatisch, wo sich Darmpolypen befinden. Der Arzt kann die weitere Behandlung gemäß des Vorschlags des Systems durchführen. Der Arzt kann auch eine eigene, vom Systemvorschlag abweichende Behandlung durchführen.</p>	<p>Zur Darmkrebs-Vorsorge wird unter anderem eine Darmspiegelung des Dickdarms durchgeführt. Hierzu führt der Arzt ein Koloskop, einen biegsamen Schlauch mit einer eingebauten Kamera, in den Darm des Patienten ein. Der Arzt kann so über einen Monitor sehen, ob die Darmschleimhaut gesund aussieht. Er vergleicht hierzu die Monitoransicht mit Normwerten. Er kann so auch tumorverdächtige Polypen erkennen und diese entfernen oder Gewebeprobe entnehmen.</p> <p>Neben dieser Untersuchungsart gibt es Koloskopie-Systeme, die auf künstliche Intelligenz (KI) zurückgreifen: Diese Systeme können selbstständig Darmpolypen über eine Bilderkennung erkennen. Die Untersuchung findet dabei weiterhin durch den Arzt statt. Verwendet wird weiterhin ein Koloskop mit einer eingebauten Kamera. Die künstliche Intelligenz greift im Hintergrund auf eine große klinische Datenbank zurück: Sie vergleicht die Daten des Patienten mit den Daten anderer Patienten. Durch diesen Datenabgleich erkennt das System auffällige Gewebeareale und markiert sie farblich.</p> <p>Das Koloskopie-System zeigt dem Arzt bei der Untersuchung auf dem Monitor automatisch, wo sich Darmpolypen befinden und gibt dem Arzt verbindlich die weitere Behandlung vor.</p>

2.2 Diagnostik: Erkennen von Herz-Kreislauf-Erkrankungen

<p>Explainability low</p>	<p>Zur Diagnose von Herz-Kreislauf-Erkrankungen wird unter anderem ein EKG (Elektrokardiogramm) durchgeführt. Bei einem EKG untersucht der Arzt das Herz des Patienten. Hierzu werden selbsthaftende Elektroden auf der Haut des Patienten befestigt. Die Elektroden erfassen die elektrische Erregung an der Hautoberfläche. Diese wird auf einem Monitor als Wellenlinie dargestellt. Der Arzt kann so über einen Abgleich mit Normwerten die Herzfrequenz und den Herzrhythmus des Patienten bewerten und Herzkrankheiten feststellen.</p> <p>Neben dieser Untersuchungsart beim Arzt gibt es EKG-Systeme, die auf künstliche Intelligenz (KI) zurückgreifen: Diese Systeme können selbstständig Messungen über eine Biosignal- und Vitalwertanalyse vornehmen. So können sie Unregelmäßigkeiten im Herz-Kreislauf-System erkennen. Das EKG-System informiert den Patienten automatisch, wenn ein Messwert außerhalb der Norm liegt. Es bietet an, zur Sicherung der Diagnose weitere Messungen vorzunehmen, die an den behandelnden Arzt weitergeleitet werden können. Der Patient kann den Vorschlag des Systems annehmen. Der Patient kann die weitere Behandlung abweichend vom Systemvorschlag auch direkt beim Arzt durchführen lassen.</p>	<p>Zur Diagnose von Herz-Kreislauf-Erkrankungen wird unter anderem ein EKG (Elektrokardiogramm) durchgeführt. Bei einem EKG untersucht der Arzt das Herz des Patienten. Hierzu werden selbsthaftende Elektroden auf der Haut des Patienten befestigt. Die Elektroden erfassen die elektrische Erregung an der Hautoberfläche. Diese wird auf einem Monitor als Wellenlinie dargestellt. Der Arzt kann so über einen Abgleich mit Normwerten die Herzfrequenz und den Herzrhythmus des Patienten bewerten und Herzkrankheiten feststellen.</p> <p>Neben dieser Untersuchungsart beim Arzt gibt es EKG-Systeme, die auf künstliche Intelligenz (KI) zurückgreifen: Diese Systeme können selbstständig Messungen über eine Biosignal- und Vitalwertanalyse vornehmen. So können sie Unregelmäßigkeiten im Herz-Kreislauf-System erkennen. Das EKG-System informiert den Patienten automatisch, wenn ein Messwert außerhalb der Norm liegt. Es nimmt weitere Messungen vor und erstellt eine verbindliche Diagnose, die an den behandelnden Arzt weitergeleitet wird.</p>
<p>Explainability high</p>	<p>Zur Diagnose von Herz-Kreislauf-Erkrankungen wird unter anderem ein EKG (Elektrokardiogramm) durchgeführt. Bei einem EKG untersucht der Arzt das Herz des Patienten. Hierzu werden selbsthaftende Elektroden auf der Haut des Patienten befestigt. Die Elektroden erfassen die elektrische Erregung an der Hautoberfläche. Diese wird auf einem Monitor als Wellenlinie dargestellt. Der Arzt kann so über einen Abgleich mit Normwerten die Herzfrequenz und den Herzrhythmus des Patienten bewerten und Herzkrankheiten feststellen.</p> <p>Neben dieser Untersuchungsart beim Arzt gibt es EKG-Systeme, die auf künstliche Intelligenz (KI) zurückgreifen: Diese Systeme können selbstständig Messungen über eine Biosignal- und Vitalwertanalyse vornehmen. So können sie Unregelmäßigkeiten im Herz-Kreislauf-System erkennen. Die Untersuchung findet dabei durch den Patienten selbst statt. Verwendet werden beispielsweise Armbanduhren, T-Shirts oder Pflaster mit eingebauten Elektroden. Diese können mit einer Smartphone-App verbunden sein. Die künstliche Intelligenz greift im Hintergrund auf eine große klinische Datenbank zurück: Sie vergleicht die Daten des Patienten mit den Daten anderer Patienten. Durch diesen Datenabgleich erkennt das System auffällige Unregelmäßigkeiten im Herz-Kreislauf-System und warnt den Patienten.</p> <p>Das EKG-System informiert den Patienten automatisch, wenn ein Messwert außerhalb der Norm liegt. Es bietet an, zur Sicherung der Diagnose weitere Messungen vorzunehmen, die an den behandelnden Arzt weitergeleitet werden können. Der Patient kann den Vorschlag des Systems annehmen. Der Patient kann die weitere Behandlung abweichend vom Systemvorschlag auch direkt beim Arzt durchführen lassen.</p>	<p>Zur Diagnose von Herz-Kreislauf-Erkrankungen wird unter anderem ein EKG (Elektrokardiogramm) durchgeführt. Bei einem EKG untersucht der Arzt das Herz des Patienten. Hierzu werden selbsthaftende Elektroden auf der Haut des Patienten befestigt. Die Elektroden erfassen die elektrische Erregung an der Hautoberfläche. Diese wird auf einem Monitor als Wellenlinie dargestellt. Der Arzt kann so über einen Abgleich mit Normwerten die Herzfrequenz und den Herzrhythmus des Patienten bewerten und Herzkrankheiten feststellen.</p> <p>Neben dieser Untersuchungsart beim Arzt gibt es EKG-Systeme, die auf künstliche Intelligenz (KI) zurückgreifen: Diese Systeme können selbstständig Messungen über eine Biosignal- und Vitalwertanalyse vornehmen. So können sie Unregelmäßigkeiten im Herz-Kreislauf-System erkennen. Die Untersuchung findet dabei durch den Patienten selbst statt. Verwendet werden beispielsweise Armbanduhren, T-Shirts oder Pflaster mit eingebauten Elektroden. Diese können mit einer Smartphone-App verbunden sein. Die künstliche Intelligenz greift im Hintergrund auf eine große klinische Datenbank zurück: Sie vergleicht die Daten des Patienten mit den Daten anderer Patienten. Durch diesen Datenabgleich erkennt das System auffällige Unregelmäßigkeiten im Herz-Kreislauf-System und warnt den Patienten.</p> <p>Das EKG-System informiert den Patienten automatisch, wenn ein Messwert außerhalb der Norm liegt. Es nimmt weitere Messungen vor und erstellt eine verbindliche Diagnose, die an den behandelnden Arzt weitergeleitet wird.</p>

2.3 Therapie: Erkennen von Über- und Unterzuckerung

	AI completeness low	AI completeness high
Explainability low	<p>Zur Therapie von Diabetes Typ 1 wird unter anderem Insulin eingesetzt. Zur Insulingabe verwenden Diabetes Typ 1-Patienten häufig einen Insulinpen im Format eines Füllfederhalters. In Abhängigkeit vom vorher gemessenen Zuckerwert stellt der Patient seine benötigte Insulinmenge am Insulinpen individuell ein. Er vergleicht hierzu seinen Insulinwert mit einem Normwert. Dann spritzt er sich die eingestellte Insulinmenge mit dem Insulinpen in die Haut. So kann der Patient seinen Zuckerhaushalt regulieren und eine Über- oder Unterzuckerung vermeiden.</p> <p>Neben dieser Behandlungsart gibt es Insulin-Systeme, die auf künstliche Intelligenz (KI) zurückgreifen: Diese Systeme können selbstständig Messungen über eine Stoffwechsel- und Glukoseanalyse vornehmen. So können sie die benötigte Insulinmenge berechnen.</p> <p>Das Insulin-System zeigt dem Patienten automatisch, wie viel Insulin er benötigt. Der Patient kann den Vorschlag des Systems annehmen und sich die angezeigte Menge Insulin verabreichen lassen. Der Patient kann seine benötigte Insulinmenge aber auch selbst bestimmen und sich eine vom Systemvorschlag abweichende Insulinmenge verabreichen lassen.</p>	<p>Zur Therapie von Diabetes Typ 1 wird unter anderem Insulin eingesetzt. Zur Insulingabe verwenden Diabetes Typ 1-Patienten häufig einen Insulinpen im Format eines Füllfederhalters. In Abhängigkeit vom Zuckerwert stellt der Patient seine benötigte Insulinmenge am Insulinpen individuell ein. Dann spritzt er sich die eingestellte Insulinmenge mit dem Insulinpen in die Haut. Er vergleicht hierzu seinen Insulinwert mit einem Normwert. So kann der Patient seinen Zuckerhaushalt regulieren und eine Über- oder Unterzuckerung vermeiden.</p> <p>Neben dieser Behandlungsart gibt es Insulin-Systeme, die auf künstliche Intelligenz (KI) zurückgreifen: Diese Systeme können selbstständig Messungen über eine Stoffwechsel- und Glukoseanalyse vornehmen. So können sie die benötigte Insulinmenge berechnen.</p> <p>Das Insulin-System zeigt dem Patienten automatisch, wie viel Insulin er benötigt und verabreicht verbindlich die benötigte Menge.</p>
Explainability high	<p>Zur Therapie von Diabetes Typ 1 wird unter anderem Insulin eingesetzt. Zur Insulingabe verwenden Diabetes Typ 1-Patienten häufig einen Insulinpen im Format eines Füllfederhalters. In Abhängigkeit vom Zuckerwert stellt der Patient seine benötigte Insulinmenge am Insulinpen individuell ein. Er vergleicht hierzu seinen Insulinwert mit einem Normwert. Dann spritzt er sich die eingestellte Insulinmenge mit dem Insulinpen in die Haut. So kann der Patient seinen Zuckerhaushalt regulieren und eine Über- oder Unterzuckerung vermeiden.</p> <p>Neben dieser Behandlungsart gibt es Insulin-Systeme, die auf künstliche Intelligenz (KI) zurückgreifen: Diese Systeme können selbstständig Messungen über eine Stoffwechsel- und Glukoseanalyse vornehmen. So können sie die benötigte Insulinmenge berechnen.</p> <p>Die Behandlung findet dabei weiterhin durch den Patienten selbst statt. Verwendet wird beispielsweise eine Insulinpumpe mit einem Glukosesensor. Diese können mit einer Smartphone-App verbunden sein. Die künstliche Intelligenz greift im Hintergrund auf eine große klinische Datenbank zurück: Sie vergleicht die Daten des Patienten mit den Daten anderer Patienten. Durch diesen Datenabgleich kann das System die benötigte Insulinmenge bestimmen.</p> <p>Das Insulin-System zeigt dem Patienten automatisch, wie viel Insulin er benötigt. Der Patient kann den Vorschlag des Systems annehmen und sich die angezeigte Menge Insulin verabreichen lassen. Der Patient kann seine benötigte Insulinmenge aber auch selbst bestimmen und sich eine vom Systemvorschlag abweichende Insulinmenge verabreichen lassen.</p>	<p>Zur Therapie von Diabetes Typ 1 wird unter anderem Insulin eingesetzt. Zur Insulingabe verwenden Diabetes Typ 1-Patienten häufig einen Insulinpen im Format eines Füllfederhalters. In Abhängigkeit vom Zuckerwert stellt der Patient seine benötigte Insulinmenge am Insulinpen individuell ein. Er vergleicht hierzu seinen Insulinwert mit einem Normwert. Dann spritzt er sich die eingestellte Insulinmenge mit dem Insulinpen in die Haut. So kann der Patient seinen Zuckerhaushalt regulieren und eine Über- oder Unterzuckerung vermeiden.</p> <p>Neben dieser Behandlungsart gibt es Insulin-Systeme, die auf künstliche Intelligenz (KI) zurückgreifen: Diese Systeme können selbstständig Messungen über eine Stoffwechsel- und Glukoseanalyse vornehmen. So können sie die benötigte Insulinmenge berechnen.</p> <p>Die Behandlung findet dabei weiterhin durch den Patienten selbst statt. Verwendet wird beispielsweise eine Insulinpumpe mit einem Glukosesensor. Diese können mit einer Smartphone-App verbunden sein. Die künstliche Intelligenz greift im Hintergrund auf eine große klinische Datenbank zurück: Sie vergleicht die Daten des Patienten mit den Daten anderer Patienten. Durch diesen Datenabgleich kann das System die benötigte Insulinmenge bestimmen.</p> <p>Das Insulin-System zeigt dem Patienten automatisch, wie viel Insulin er benötigt und verabreicht verbindlich die benötigte Menge.</p>

KI-BA:

**Künstliche Intelligenz in der Versorgung – Bedingungen der
Akzeptanz von Versicherten und ÄrztInnen**

Studiendesign

21.02.2023

Stand



Unterschrift

Dr. Stefan T. Kamin

Verantwortliche Person für die Methodik

Inhalt

STUDIENDESIGN	3
KURZBESCHREIBUNG	3
KATEGORISIERUNG	3
STICHPROBENDESIGN VERSICHERTE.....	3
FALLZAHLBERECHNUNG VERSICHERTE	4
STICHPROBENDESIGN ÄRZTINNEN	4
FALLZAHLBERECHNUNG ÄRZTINNEN	4
AUSWERTUNG	5
VERSUCHSPLAN	5
AUFBAU DER BEFRAGUNG	5

Studiendesign

Kurzbeschreibung

Grundanliegen des Vorhabens ist es zu verstehen, wie die Akzeptanz gegenüber KI-Systemen bei Versicherten und ÄrztInnen ausgeprägt ist und welche Einflussfaktoren die individuelle Akzeptanz beeinflussen. Dafür werden den Probandinnen und Probanden im Rahmen eines 2x2x3-faktoriellen Designs randomisierte KI-Anwendungsszenarien in Form animierter Kurzvideos präsentiert. Die Videos repräsentieren die zuvor entwickelte Kategorisierung und sind mit Blick auf die individuelle Akzeptanz gegenüber der entsprechenden KI-Anwendung zu bewerten.

Kategorisierung

In einem ersten Schritt wird auf Grundlage eines systematischen Reviews und Experteninterviews ein Kategorisierungsinstrument für KI-Anwendungen in der medizinischen Versorgung erstellt. Das mehrdimensionale Kategorisierungsinstrument soll der interdisziplinären Sicht (Sicht von Patienten, Ärzten, Fachdisziplinen, Versorgungssektoren, etc.) gerecht werden. Anhand der Kategorisierung werden die Stimuli, stellvertretend für exemplarische Vertreter von KI-Systemen für die Befragung entwickelt, die jeweils eine typische Kombination von Merkmalsausprägungen der Dimensionen haben. Die künstlichen Stimuli beinhalten die KI-Methoden, das Einsatzgebiet, Datenheterogenität und beteiligten Akteure und sind als animierte Video-Kurzsequenzen umgesetzt. Die Entwicklung der Videos erfolgt mit den Domänenexperten und unter Zuhilfenahme des Kategorisierungsinstruments. Im Ergebnis existieren archetypische Vertreter jeder Kategorie, die schnell von Probanden rezipiert und verstanden werden können. Um mögliche Verzerrungen durch das Design der künstlichen Stimuli ausschließen zu können, werden Pretests mit ÄrztInnen (n = 5) und Versicherten (n = 5) durchgeführt. In einer weiteren Iteration werden die Rückmeldungen aus den Pretests inhaltsanalytisch ausgewertet und zur Optimierung der Stimuli herangezogen.

Stichprobendesign Versicherte

Die erstellten Video-Stimuli werden im Rahmen einer querschnittlichen Akzeptanzstudie durch Versicherte und ÄrztInnen bewertet. Die Ziehung der Versicherten-Stichprobe (N= 1.536 Personen über 18 Jahre in Privathaushalten) erfolgt durch eine mehrstufige Zufallsauswahl auf Basis lokaler Melderegister in Nordbayern. In der ersten Stufe werden 50 Gemeinden zufällig und proportional zur Bevölkerungszahl ausgewählt. Dieses Vorgehen gewährleistet eine breite regionale Repräsentanz der Teilnehmenden und stellt zugleich eine akzeptable Balance zwischen Klumpungsfaktor und Kostenaufwand sicher. Die zweite Stufe umfasst die Ziehung einer disproportional nach Alter geschichteten Zufallsstichprobe. Dazu werden aus den Auswahleinheiten (Gemeinden) der ersten Stufe gleich große Zufallsstichproben aus drei Altersgruppen gezogen (18-39 Jahre, 40-59 Jahre, 60 Jahre und älter). Dieses Vorgehen wurde gewählt, um extreme Zufallsstichproben (z.B. Überrepräsentation jüngerer Teilnehmer) zu vermeiden und sicherzustellen, dass Teilnehmende unterschiedlicher Altersgruppen vertreten sind. Um in die Studie eingeschlossen zu werden, müssen Personen über 18 Jahre alt sein. Nicht eingeschlossen werden Personen, die (dauerhaft) in stationären Einrichtungen leben. Die gezogenen Personen werden postalisch kontaktiert und zur Teilnahme an der CAWI-Befragung eingeladen. Personen, die über keinen Internetzugang verfügen, erhalten eine CAPI-Befragung. Diese wird in der Wohnung der jeweiligen Person durchgeführt. Bei Nichterreichung der

Fallzahlen erfolgt eine Ausweitung der Rekrutierung in Form einer online-basierten Ad-hoc Stichprobe auf das gesamte Bundesgebiet. Dabei ist die Repräsentativität der ursprünglich auszuwertenden Region Nordbayern weiterhin gegeben. Eventuelle Selektionseffekte, die sich durch die Ausweitung der Rekrutierung ergeben könnten, werden statistisch kontrolliert und im Rahmen der Projektergebnisse kritisch reflektiert. Die Stichproben können jederzeit unterschieden werden; zudem kann durch Gewichtungsverfahren eine Anpassung an bekannte Randverteilungen der vorgesehenen Stichprobe erfolgen.

Fallzahlberechnung Versicherte

Für die Fallzahlberechnung werden zwei zentrale Parameter zu Grunde gelegt: (1) ein angestrebtes Konfidenzniveau von 95% und (2) eine tolerierte Fehlerwahrscheinlichkeit von 2.5%. Das Konfidenzintervall (Z) beschreibt die Präzision der Parameterschätzung innerhalb der Population und wird im Folgenden mit einem üblichen Niveau von 95% für die Fallzahlberechnung herangezogen. Die Fehlerwahrscheinlichkeit (E) beschreibt die Abweichung der Stichprobe von den Werten in der Grundgesamtheit und wird im Folgenden mit einem Niveau von 2.5% in die Berechnung einbezogen. Die Berechnung erfolgt nach der Formel $n = p(100-p)z^2/E^2$ und ergibt eine anzustrebende Gesamtstichprobe von N = 1.536 Personen (n = 512 Personen je Altersgruppe). Die Berechnung der Teststärke erfolgt anhand des unten beschriebenen Versuchsplans in einem 2x2x3-faktoriellen Design. Für die Berechnung werden entsprechend 12 Gruppen sowie 10 Kovariaten als Berechnungsgrundlage angenommen. In diesem Rechenmodell ist ein Stichprobenumfang von N = 1.536 Personen ausreichend, um einen kleinen Effekt ($f = .10$) mit einer Wahrscheinlichkeit von 77% zu entdecken.

Stichprobendesign ÄrztInnen

Die Grundgesamt der ÄrztInnen-Stichprobe umfasst das ambulante Versorgungsangebot der hausärztlichen und fachärztlichen Versorgung sowie in Krankenhäusern und Kliniken angestellte ÄrztInnen in Nordbayern. Bei ÄrztInnen handelt es sich um eine Subpopulation, für die klassische Zufallsverfahren nur mit hohem Aufwand zu realisieren sind. Aus diesem Grund folgt die Ziehung dieser Stichprobe einem bewussten Auswahlverfahren durch direkte Ansprache unterschiedlicher Akteure, um eine möglichst heterogene Stichprobe im Hinblick auf die Fachdisziplin, das Versorgungsumfeld, die Region und die Berufserfahrung abzubilden. Der angestrebte Stichprobenumfang für die ÄrztInnen liegt bei N= 475 Personen. Bei Nichterreichung der Fallzahlen erfolgt eine Ausweitung der Rekrutierung auf das gesamte Bundesgebiet. Die Repräsentativität der ursprünglich auszuwertenden Region Nordbayern ist auch durch die Ausweitung der Rekrutierung weiterhin gegeben. Ebenso wird durch die Ausweitung der Befragungsregion keine Auswirkung auf die Auswertung des räumlichen Kontexts erwartet: Der räumliche Kontext als einer von mehreren potenziellen Einflussfaktoren der KI-Akzeptanz wird im vorgelegten Studiendesign über ein Item zur Stadt- bzw. Gemeindegröße des Arbeitsorts erfasst, das eine Evaluation des räumlichen Kontexts unabhängig von regionenspezifischen Besonderheiten erlaubt. Ferner ist durch die Abfrage der Postleitzahl stets eine Filtermöglichkeit gegeben.

Fallzahlberechnung ÄrztInnen

Für die Fallzahlberechnung werden zwei zentrale Parameter zu Grunde gelegt: (1) ein angestrebtes Konfidenzniveau von 95% und (2) eine tolerierte Fehlerwahrscheinlichkeit von 4,5%. Das Konfidenzintervall (Z) beschreibt die Präzision der Parameterschätzung innerhalb der Population und wird im Folgenden mit einem üblichen Niveau von 95% für die

Fallzahlberechnung herangezogen. Die Fehlerwahrscheinlichkeit (E) beschreibt die Abweichung der Stichprobe von den Werten in der Grundgesamtheit und wird im Folgenden mit einem Niveau von 4,5% in die Berechnung einbezogen. Aufgrund des geringeren avisierten Stichprobenumfangs fällt die Fehlerwahrscheinlichkeit im Vergleich zur Versichertenstichprobe höher aus. Allerdings liegt diese mit unter 5% in einem akzeptablen und üblichen Bereich. Die Berechnung erfolgt nach der Formel $n = p(100-p)z^2/E^2$ und ergibt eine anzustrebende Gesamtstichprobe von $N = 475$ Personen. Die Berechnung der Teststärke erfolgt anhand des unten beschriebenen Versuchsplans in einem 2x2x3-faktoriellen Design. Für die Berechnung werden entsprechend 12 Gruppen sowie 10 Kovariaten als Berechnungsgrundlage angenommen. In diesem Rechenmodell ist ein Stichprobenumfang von $N = 475$ Personen ausreichend, um einen mittleren Effekt ($f = .25$) mit einer Wahrscheinlichkeit von 98% zu entdecken.

Auswertung

Die KI-Akzeptanz ist als latentes multidimensionales Konstrukt konzeptualisiert. Aus diesem Grund erfolgt die Auswertung mittels latenter Strukturgleichungsmodelle. Neben der Analyse der Zusammenhänge zwischen Einflussfaktoren und KI-Akzeptanz erlauben Strukturgleichungsmodelle auch den Vergleich zwischen verschiedenen Anwendungsszenarien und Probandengruppen. Durch sogenannte Multigruppenmodelle können simultan die Daten von Versicherten und Ärztinnen analysiert werden, um Unterschiede und Gemeinsamkeiten zwischen beiden Gruppen in Bezug auf die Parameterschätzungen zu identifizieren. Dabei wird auch das im Versuchsplan abgebildete 2x2x3-faktorielle Design entlang der unterschiedlichen Bedingungen geprüft. Durch dieses Vorgehen kann sichtbar gemacht werden, inwiefern bestimmte Zusammenhänge in verschiedenen Populationen gleich oder unterschiedlich ausfallen. Weiterhin ermöglichen Mehrgruppenvergleiche die Bewertung der psychometrischen Qualität der Messungen. So kann beispielsweise sichergestellt werden, dass die Erfassung der KI-Akzeptanz äquivalent zwischen Versicherten und ÄrztInnen erfolgt und Unterschiede zwischen Parameterschätzungen nicht auf Messfehler zurückzuführen sind.

Versuchsplan

Der mehrfaktorielle Versuchsplan beinhaltet zwei unabhängige Variablen mit jeweils zwei Faktorstufen und eine unabhängige Variable mit 3 Faktorstufen (2x2x3 Design):

- IV1: AI-Explainability (high/low), d.h. die Funktionsweise der KI wird dargestellt – oder nicht
- IV2: AI-Completeness (high/low), d.h. die KI entscheidet autonom – oder der Mensch
- IV3: KI Use-case (Diabetes, EKG, Koloskopie)

Aufbau der Befragung

Der Aufbau der Befragung ist in Abb. 1 dargestellt. Die erste Seite umfasst die Erhebung grundlegender soziodemographischer Informationen. Danach werden die Probanden gebeten, sich ein Video zur Einführung in die KI-Thematik anzuschauen. Nach diesem Video erfolgt die Erhebung der KI-Akzeptanzskala, die im Rahmen des Projektes entwickelt wurde, um interindividuelle Unterschiede in der Bewertung der KI-Lösungen zu erklären. Danach erfolgt die randomisierte Zuteilung zur einer between-subject Bedingung, innerhalb derer eine randomisierte Videosequenz zu evaluieren ist. Die Evaluation erfolgt anhand klassischer Instrumente aus der Technikakzeptanz- sowie der UX-Forschung, die gleichzeitig als zentrale

abhängige Variablen fungieren. Nach der Evaluationen werden die Probanden gebeten, weitere Fragen zu potentiellen Einflussfaktoren der KI-Akzeptanz zu beantworten. Dabei handelt es sich um die Bereiche:

- Räumlicher Kontext (z.B. Versorgungssituation, Gemeindegröße)
- Soziale Beziehungen (z.B. Anzahl Beziehungspartner, soziale Unterstützung)
- Gesundheitliche Ressourcen (z.B. Krankheiten, Alltagsfunktionalität)
- Gesundheitsverhalten (z.B. Vorsorgeverhalten, Arzt-Patienten-Beziehung)
- Persönlichkeit (z.B. Zeitperspektive, Alterserleben)
- Technik und Digitalisierung (z.B. Technikmotivation, Digitalkompetenz)

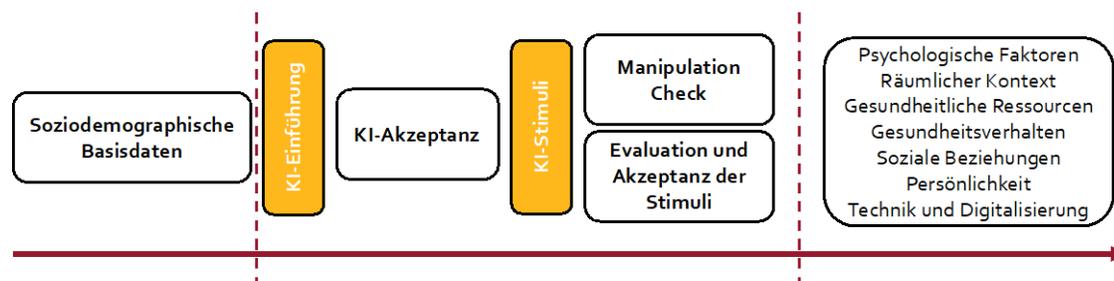


Abbildung 1: Aufbau der Befragung

Ergebnisdokumentation

01VSF20016 - KI-BA
KÜNSTLICHE INTELLIGENZ IN DER VERSORGUNG
- BEDINGUNG DER AKZEPTANZ VON
VERSICHERTEN UND ÄRZTINNEN

AP 9 „Handlungsempfehlungen“

Inhalt

1	Einleitung	3
2	Vorgehensweise	4
2.1	Durchführung des Expertenworkshops	4
2.2	Auswertung des Kurzfragebogens	5
2.3	Auswertung des Expertenworkshops	5
3	Ergebnisse	7
3.1	Ergebnisse des Kurzfragebogens und Beschreibung der Stichprobe	7
3.2	Ergebnisse des Expertenworkshops.....	8
3.2.1	Grundlegende Definition von KI.....	8
3.2.2	Grundlegende Rahmenbedingungen	10
3.2.3	Entwicklung: Akzeptanzobjekt	11
3.2.4	Implementierung: Akzeptanzkontext und Akzeptanzsubjekt	14
3.2.4.1	Akzeptanzkontext	14
3.2.4.2	Bezogen auf Akzeptanzsubjekt	16
3.3	Ergebnisse der Priorisierung	18
4	Zusammenführung	20
5	Fazit	23
6	Literatur	24
7	Anhang	25
7.1	Workshop-Dokumentation	25
7.2	MAXQDA-Codesystem	26
Abbildungsverzeichnis		29
Tabellenverzeichnis		30
1	Einleitung	4
2	Vorgehensweise	5
2.1	Durchführung des Expertenworkshops	5
2.2	Auswertung des Kurzfragebogens	6
2.3	Auswertung des Expertenworkshops	6
3	Ergebnisse	8
3.1	Ergebnisse des Kurzfragebogens und Beschreibung der Stichprobe	8
3.2	Ergebnisse des Expertenworkshops.....	9
3.2.1	Grundlegende Definition von KI.....	9
3.2.2	Grundlegende Rahmenbedingungen	11
3.2.3	Entwicklung: Akzeptanzobjekt	12
3.2.4	Implementierung: Akzeptanzkontext und Akzeptanzsubjekt	15
3.2.4.1	Akzeptanzkontext	15
3.2.4.2	Bezogen auf Akzeptanzsubjekt	17
3.3	Ergebnisse der Priorisierung	19
4	Zusammenführung	21
5	Fazit	24
6	Literatur	25
7	Anhang	26
7.1	Workshop-Dokumentation	26

Anhang 5: Ergebnisdokumentation AP 9 – Handlungsempfehlungen

7.2 MAXQDA-Codesystem27

Abbildungsverzeichnis 30

Tabellenverzeichnis 31

1 Einleitung

Bedingungen für die Akzeptanz von KI in der Gesundheitsversorgung können vielfältig sein. Um KI Anwendungen dauerhaft erfolgreich in der Praxis zu implementieren, ist es wichtig zu verstehen, welche Bedingungen die Akzeptanz von KI in der Gesundheitsversorgung beeinflussen. Aus diesem Grund untersucht das Projekt KI-BA die Akzeptanz von KI-Anwendungen in beispielhaften Versorgungsbereichen der GKV und verfolgt dabei insgesamt drei Projektziele. Projektziel 1 beinhaltet die **repräsentative Auswahl praxisnaher KI-Anwendungsszenarien** mit hohem Potential zur Entlastung der Akteure¹ (Kategorisierung). Projektziel 2 umfasst die **Durchführung einer Akzeptanzstudie**, in deren Rahmen die entwickelten KI-Anwendungsszenarien durch Versicherte und Ärzte evaluiert und bewertet werden. Die Befragung von Versicherten und Ärzte erlaubt erstmals eine multiperspektivische Identifikation von individuellen Anforderungen, Herausforderungen und Barrieren, die mit der (direkten oder indirekten) Nutzung von KI-Anwendungen in der Versorgung einhergehen. Ziel der letzten Projektphase ist die **Translation der Studienergebnisse in konkrete Handlungsempfehlungen für den Einsatz von KI in der Versorgung**, um die Ergebnisse der Akzeptanzstudie – neben der wissenschaftlichen Veröffentlichung – auch für Akteure aus der Praxis nutzbar zu machen. Im Rahmen eines Expertenworkshops wurden hierzu zum einen **Implikationen für die nutzerzentrierte Entwicklung und Gestaltung von KI-Systemen**, zum anderen **Strategien zur Implementierung von KI-Systemen in der GKV** erarbeitet und in **konkrete Handlungsempfehlungen** überführt. Der nachfolgende Ergebnisbericht schildert zeigt die Vorgehensweise und Ergebnisse aus der letzten Projektphase: Die Durchführung des Expertenworkshops mit dem Ziel der Entwicklung von Handlungsempfehlungen für die Gestaltung und Implementierung von KI-Systemen in der Gesundheitsversorgung.

Ergebnisse: Im Zuge der Auswertung konnten nicht nur wie ursprünglich geplant Handlungsempfehlungen entworfen werden, sondern konkrete Anforderungen, welche eine KI erfüllen sollte herausgearbeitet werden. Anforderungen wurden in vier Themenbereichen identifiziert. Vor der Entwicklung und Implementierung muss sowohl die KI, als auch deren Rahmenbedingungen grundlegend definiert werden. Zudem können Anforderungen zwischen allgemein zutreffenden Anforderungen und spezifisch für Ärzte zutreffenden Anforderungen entschieden werden. Die Berücksichtigung dieser Punkte wirkt sich letztlich positiv auf die Akzeptanz einer KI aus. Zudem konnte eine erste Antwort auf die Forderung nach einer Definition von KI herausgearbeitet werden.

¹ Aufgrund der besseren Lesbarkeit wird im Text das generische Maskulinum verwendet. Gemeint sind jedoch immer alle Geschlechter.

2 Vorgehensweise

Um **Implikationen für die nutzerzentrierte Entwicklung und Gestaltung von KI-Systemen**, sowie **Strategien zur Implementierung von KI-Systemen in der Gesundheitsversorgung** zu erarbeiten und in **konkrete Handlungsempfehlungen** zu überführen wurde ein Expertenworkshop durchgeführt. Ziel des Workshops war es, den identifizierten Akzeptanzrisiken und Barrieren bei der Einführung von KI-Technologien zu begegnen. Um diesen Workshop-Zielen gerecht werden zu können, wurde die Zusammensetzung einer möglichst heterogenen und repräsentativen Teilnehmergruppe, angestrebt. Die Rekrutierung erfolgte dabei a) über eine Akquise im Rahmen der in der Kategorisierungsphase erfolgten Experteninterviews, b) über Ansprechpartner der jeweiligen Krankenkassen und c) über eine direkte persönliche Ansprache. Da es sich beim Expertenworkshop um ein qualitatives Verfahren handelt, war keine Zufallsauswahl der Stichprobe nötig, vielmehr wurde bei der Auswahl der Experten darauf geachtet, dass sie für die Workshop-Ziele möglichst eine Vielfalt an fachlichem Input liefern können (Tausch & Menold, 2015). Im Zuge der Auswertung konnten nicht nur Handlungsempfehlungen entworfen, sondern konkrete Anforderungen an die KI entworfen werden, weshalb im Folgenden nicht mehr von Handlungsempfehlungen, sondern von Anforderungen gesprochen wird.

2.1 Durchführung des Expertenworkshops

Der Workshop wurde auf Basis eines **Fokusgruppen-Konzepts** durchgeführt und folgt einer **methodischen Trias** (Hackel & Klebl, 2008): In einem ersten Schritt wurde der **gesamten Expertenrunde die Ergebnisse der Akzeptanzstudie vorgestellt**. Im nächsten Schritt **erarbeiteten die Experten in moderierten Kleingruppen mittels Design-Thinking-Methoden eine Sammlung konkret in die Praxis umsetzbarer Implikationen**. Im letzten Schritt der methodischen Trias folgte eine **konsolidierende Diskussion in der Gesamtgruppe**, woraus eine erste grobe Sammlung praktischer Anforderungen entstand.

Für den Expertenworkshop wurde eine Dauer von ca. 5 Stunden inkl. Pausen angesetzt.

Im Anschluss an die Begrüßung und einem ersten Kennenlernen folgte das Ausfüllen des Kurzfragebogens und der Einverständniserklärung. Zur Beschreibung der Stichprobe wurden folgende Aspekte abgefragt:

- Geschlecht (männlich, weiblich, divers)
- Alter
- Beruf
- Dauer der Beschäftigung
- Technikbereitschaft

Einschätzungen zur Technikbereitschaft der Befragten wurden auf einer 5-Punkt-Likert-Skala (1=stimmt gar nicht, 5=stimmt völlig) abgefragt. Die Kurzskaala basiert auf einem Modell der Technikbereitschaft, welches drei

unterscheidbare Facetten als Determinanten für individuell unterschiedliche Bereitschaft zum Umgang mit Technik konstatiert: Technikakzeptanz, Technikkompetenz- und Technikkontrollüberzeugungen (Neyer, Felber & Gebhardt, 2016). Insgesamt wurden 12 Items abgefragt, die einen Einblick in das Erleben und die Einstellung der Befragten bezüglich der Anwendung neuer Technologien/Elektronik geben sollten.

Zur anschließenden Diskussion in moderierten Kleingruppen wurde den Teilnehmenden verschiedene Triggerbegriffe als Anhaltspunkt vorgelegt. Diese wurden in Anlehnung an Lucke (1995) gewählt. Berücksichtigt wurden dabei die drei Konzepte Akzeptanzobjekt, Akzeptanzsubjekt und Akzeptanzkontext. Diese wurden zum einfacheren Verständnis durch die Begriffe Technologie, Individuum und Nutzungskontext ersetzt. Wie bereits erwähnt wurden besonders zwei Teilbereiche berücksichtigt. Dabei handelt es sich zu einem um die Gestaltung und Entwicklung, zum anderen um die nutzerzentrierte Implementierung.

Übersicht über die Triggerbegriffe für die Priorisierung:

- Nutzerzentrierte Gestaltung und Entwicklung:
 - Technologie
 - Individuum
 - Nutzungskontext
- Nutzerzentrierte Implementierung:
 - Technologie
 - Individuum
 - Nutzungskontext

In einer abschließenden Gruppendiskussion wurden die Teilnehmer gebeten als von zentraler Bedeutung empfundene Aspekte auf Post-Its zu notieren (siehe 7.1). Diese ergeben eine Priorisierung der Workshop Inhalte.

2.2 Auswertung des Kurzfragebogens

Der Kurzfragebogen wurde mittels SPSS und Excel deskriptiv ausgewertet. Auswertungskriterien waren das Geschlecht, Alter, sowie die Berufserfahrung der Befragten. Zudem wurde die Technikbereitschaft der Befragten anhand der Determinanten Technikakzeptanz, Technikkompetenz- und Technikkontrollüberzeugungen ermittelt. Die Ergebnisse wurden in eine gemeinsame Tabelle überführt.

2.3 Auswertung des Expertenworkshops

Um die Ergebnisse des Expertenworkshops im Anschluss in einem Katalog mit Anforderungen an eine nutzergerechte Implementierung von KI-Systemen in der Versorgung sowie als Basis für zukünftige agile Entwicklungsprozesse von KI-Systemen konsolidieren zu können, wurden Audioaufzeichnungen des Workshops erstellt und Diskussionsinhalte schriftlich (Post-Its) und fotografisch dokumentiert (siehe 7.1). Für die Datenauswertung erfolgte eine Transkription und qualitative Inhaltsanalyse

Anhang 5: Ergebnisdokumentation AP 9 – Handlungsempfehlungen

in Anlehnung an Mayring (2015) unter Berücksichtigung der zusätzlichen Materialien.



3 Ergebnisse

Der Ergebnisteil lässt sich in drei Teilbereiche unterteilen. Zuerst werden die Ergebnisse des Kursfragebogens dargestellt, welche grundlegende Erkenntnisse über die Stichprobe des Workshops ermöglichen. Den zweiten Teil stellen die Ergebnisse der Gruppendiskussion des Workshops dar. Abschließend werden die Inhalte der Priorisierung dargestellt. Aus den im Zuge des Expertenworkshops diskutierten Chancen und Risiken wurde ein Codesystem mit verschiedenen Ober- und Unter-codes generiert, welches die Grundlage der Ableitung von Anforderungen an die KI darstellt.

3.1 Ergebnisse des Kurzfragebogens und Beschreibung der Stichprobe

Insgesamt nahmen 13 Experten am Workshop teil. Darunter ordneten sich 5 dem weiblichen und 8 dem männlichen Geschlecht zu. Die Experten kommen aus dem wissenschaftlichen und medizinischen Bereich (Ärzte, Pathologie Qualitätsmanagement, Pflege, Gerontologie, Medizintechnik). Zudem kommen die Experten aus dem Consulting und Projektmanagement.

Tabelle 1: Ergebnisse Kurzskala Technikbereitschaft

	<i>M</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>
<i>Alter</i>	42,25	9,6	25	58
<i>Berufserfahrung</i>	Weniger als 1 Jahr: 6,7% 1-3 Jahre: 13,3% 3-5 Jahre: 6,7 % mehr als 10 Jahre: 60%			
<i>Technikbereitschaft*</i>	46,6	5,5	35	54
<i>Technikakzeptanz**</i>	15,2	3,4	6	20
<i>Technikkompetenz**</i>	17,4	2,3	12	20
<i>Technikkontrollüberzeugung**</i>	14,0	2,0	11	17

* von insgesamt 60 möglichen Punkten,

** von insgesamt 20 möglichen Punkten

basierend auf Neyer, F. J., Felber, J., & Gebhardt, C. (2016). Kurzskala Technikbereitschaft (TB, technology commitment).

3.2 Ergebnisse des Expertenworkshops

Die im Zuge der Auswertung mit MAXQDA erstellten Codes stellen die Grundlage der Anforderungen dar und sind in verschiedene Themenbereichen unterteilt, welche im Folgenden dargestellt werden. Insgesamt wurden 142 Codes identifiziert (ausführliches Codesystem siehe 7.2). Die Codes wurden im Folgenden zur besseren Übersicht zu Themenbereichen aggregiert und schriftlich dargestellt. Aus der Diskussion abgeleitete Anforderungen an die KI konnten in folgende Themenbereiche unterteilt werden:

- Grundlegende Definition von KI
- Grundlegende Rahmenbedingungen
- Entwicklung
- Implementierung

Zusätzlich konnte in einem weiteren Schritt zwischen Anforderungen für verschiedene Personengruppen unterschieden werden. Berücksichtigt wurden allgemeine Anforderungen und Anforderungen, welche spezifisch für Ärzte zutreffend sind.

Im Folgenden werden die Anforderungen nach identifizierten Themenbereichen dargestellt und erläutert.

3.2.1 Grundlegende Definition von KI

Im Zuge der Auswertung konnte festgestellt werden, dass die grundlegende Definition von KI eine zu klärende Problematik darstellt. Dem Themenbereich konnten insgesamt 34 Codes zugeordnet werden. Dabei wurden 15 Codes im Bereich der allgemeinen Anforderungen identifiziert, während 19 Codes spezifisch auf Ärzte zutreffen. Erstere sind im Folgenden dargelegt.

Grundlegende Definition von KI:

Definition von "KI" notwendig

Die erste abgeleitete Anforderung deckt sich mit dem Themenbereich der Definition von KI. Besonders zu berücksichtigen ist dabei, was die KI bewirken soll. Eine im Zuge dessen genannte Anforderung ist, dass KI als kritische Hinterfragungskomponente zu verstehen sein soll. Konkret soll die KI stets eine Fehlerreduktion als Ziel verfolgen und zur Reflexion von Arbeitsschritten anregen, indem sie medizinisches Personal beispielsweise an Arbeitsschritte und to-Dos erinnert. Gleichzeitig sollte die KI eine Funktion haben, in welcher erklärt wird, warum die KI Bestimmtes vorschlägt.

Definition der Rolle des Menschen in der KI

Eine weitere Anforderung ist, dass die Rolle des Menschen im Umgang mit der KI zu definieren ist. Dies bezieht sich vor allem auf durch einen Algorithmus generierte Statistiken. Hier wurde die Frage aufgeworfen, was in diesem Kontext die Rolle des Menschen sei.

Berücksichtigung des Einflusses von KI auf das Arbeitsfeld der Medizin

Ein weiterer Themenbereich ist die Berücksichtigung des Einflusses von KI auf das Arbeitsfeld der Medizin im Allgemeinen. Darunter wurden einerseits zu berücksichtigende Aspekte, andererseits Anforderungen an die Rolle von KI entworfen. Zu berücksichtigen ist der Einfluss von KI auf die Arbeitsprozesse in der Medizin, die Entmenschlichung von KI zur Herstellung von Neutralität, sowie die Berücksichtigung des Einflusses der Visualisierung der KI auf deren Wahrnehmung. KI sollte als Empfehlung im Entscheidungsprozess, als Statistikpool und allgemein als Werkzeug angesehen werden.

Wirtschaftliche Unabhängigkeit/Neutralität

Verdeutlichung der Rolle von KI als unterstützendes Tool

Weitere Anforderungen sind die Garantie wirtschaftlicher Unabhängigkeit und Neutralität, sowie die Verdeutlichung, dass KI als unterstützendes Tool und nicht als selbstständig handelnd zu betrachten ist. Darunter zählt auch die Garantie, dass keine personenbezogenen Daten in die KI integriert werden.

Abgrenzung von KI zu bereits eingesetzter Technik

Zudem wurde die Frage aufgeworfen, inwiefern sich die hier im Projekt vorgestellte KI von bereits eingesetzten Technologien unterscheidet, was zu der Überlegung führte, ob KI überhaupt als solche gekennzeichnet werden sollte.

Als spezifisch für Ärzte zutreffend eingestuft wurden folgende Punkte:

KI in der Lehre/ zur Schulung von medizinischem Personal

Anforderungen, welche spezifisch auf Ärzte zutreffen, sind der Einsatz von KI in der Lehre und zur Schulung von medizinischem Personal. Besonders wurde die Fähigkeit der KI als „geduldiger Lehrer“ im Vergleich zum Menschen hervorgehoben. Im Zuge dieser Anforderung wird einschränkend angemerkt, dass der kreative Faktor beim Einsatz von KI in der Lehre wegfallen.

Kein Ersetzen von Ärzten durch KI

Ein wichtiger Themenbereich bildet die Anforderung, die Ersetzbarkeit von Ärzten durch KI zu vermeiden. Weiterhin wurde ausgeführt, dass sich der Einsatz von KI vor allem im pathologischen Bereich eignet. Um die Ablehnung von KI zu vermeiden wird Aufklärung und Transparenz, sowie das Hervorheben, dass KI Emotionalität und somit Ärzte nicht ersetzen kann, empfohlen. Zudem soll eine exzessive Nutzung und Auseinandersetzung mit der KI zur Vermeidung der Ersetzbarkeit von Ärzte beitragen.

Definition des Nutzens und Einsatzbereichs der KI

(Technische) Nachvollziehbarkeit für Ärzte

Weitere Anforderungen sind die Definition des Nutzens und Einsatzbereichs der KI, sowie die Ermöglichung technischer Nachvollziehbarkeit für Ärzte. Hierbei ist anzufügen, dass von den Teilnehmenden grundsätzlich eine höhere Akzeptanz von KI im diagnostischen Bereich und eine niedrigere Akzeptanz im therapeutischen Bereich angenommen wird.

KI zur Verringerung der Arbeitsbelastung von Ärzte

Eine letzte Anforderung, welche spezifisch für Ärzte zutreffend ist, ist die Anforderung, dass die KI eine Verringerung der Arbeitsbelastung für Ärzte bewirken soll. Darunter zählt die Steigerung der Zeiteffektivität und Verbesserung der Prozessstrukturen durch den Einsatz von KI, sowie die Steigerung der diagnostischen Sicherheit.

3.2.2 Grundlegende Rahmenbedingungen

Der zweite Themenbereich identifiziert grundlegende Rahmenbedingungen, welche von potentiellen KI Anwendungen erfüllt werden sollten. Von den insgesamt 16 Codes, die diesem Themenbereich zugeordnet werden konnten, entfielen 10 Codes auf allgemeine Anforderungen und 6 Codes auf Anforderungen für die Gruppe der Ärzte.

Im Folgenden sind die aus der Diskussion abgeleiteten und allgemein zutreffenden Anforderungen dargestellt:

KI "bereit" für den Menschen machen

Überwinden von bürokratischen Hürden bei Einführung von KI

Eine der Anforderungen ist, die KI „bereit“ für den Menschen zu machen. Zudem müssen bürokratische Hürden bei der Einführung von KI überwunden werden.

Ethische Anforderungen

Des Weiterem muss die KI ethische Anforderungen erfüllen. Hier wurden die Vermeidung von Bias beim Anlernen der KI, sowie die Sicherung der Datenqualität als Grundlage des Algorithmus genannt.

Standardisierung innerhalb der EU

Hinzu kommt eine Standardisierung innerhalb EU, welche gleichzeitig die Unabhängigkeit und Neutralität der KI gewährleisten soll.

Abhängigkeit des Vertrauens von Patienten vom Nutzungskontext

Zudem wurde angemerkt, dass das Vertrauen von Patienten in die KI von deren Nutzungskontext abhängt. Aus diesem Grund sind weitere Anforderungen die Berücksichtigung der übergeordneten Rolle von Kontrolle auf der Seite der Patienten, wohingegen die Einbeziehung von Patienten eine untergeordnete Rolle zukommt.

Im Folgenden sind die spezifisch für Ärzte zutreffenden Anforderungen dargestellt:

Ärzte dürfen sich nicht auf KI verlassen

Hieraus lässt sich ableiten, dass Ärzte sich nicht auf die KI verlassen dürfen. Hierfür soll eine Pflichtkomponente in die KI integriert werden, welche das autonome Arbeiten der Ärzte trotz des Einsatzes von KI unterstützt.

Klärung der Haftung und Eigenbestimmung der Ärzte

Des Weiteren muss die Haftung und Eigenbestimmung der Ärzte im Umgang mit der KI geklärt werden. Während der Fokusgruppendifferenzierung wurde sich für die Haftung von Ärzten ausgesprochen. Zudem soll eine Zielgruppendifferenzierung berücksichtigt werden. Beispielsweise wurde angenommen, dass es sowohl Ärzte gibt, die "Follower" sind und sich an Vorgaben und Leitlinien halten, als auch „Leader“, die eigenständig Entscheidungen treffen wollen. Dies sollte bei der Gestaltung einer KI berücksichtigt werden.

3.2.3 Entwicklung: Akzeptanzobjekt

Einen weiteren Fokus in der Gruppendiskussion stellt die eigentliche Entwicklung der KI dar. Hierunter fallen alle genannten Inhalte und Anforderungen, welche für die technische Entwicklung von Bedeutung sind. Hier konnte eine Unterteilung in Anlehnung an Lucke (1995) vorgenommen werden. Inhalte des Themenblocks Entwicklung decken sich mit dem oben

genannten Begriff Technologie, beziehungsweise Akzeptanzobjekt, weshalb der Begriff als Übercode eingeführt wurde. Insgesamt wurden 49 Codes der Entwicklung, beziehungsweise dem Akzeptanzobjekt zugeordnet. Auch hier fand die Unterteilung in „Allgemein“ mit 39 Codes und „spezifisch auf Ärzte zutreffend“ mit 10 Codes statt.

Im Folgenden sind die Anforderungen zum Akzeptanzobjekt und allgemein zutreffend dargestellt:

Generierung der Datenbank der KI als wichtiger Punkt

Für das Akzeptanzobjekt von Bedeutung ist die Generierung der Datenbank der KI. Diese sieht sich mit zwei Problemen konfrontiert. Darunter zum einen die Aktualität der Daten, zum anderen deren Ursprung, sowie die Validität und Güte der Daten. Im Zuge dessen genannte Anforderungen stellt die Entwicklung einer ausgereiften, validen Datenbank noch vor der Entwicklung, sowie die zentrale Bedeutung valider Daten dar.

Bevorzugung analytischer statt generativer KI

Während der Fokusgruppendifkussion stellte sich heraus, dass analytische KIs generativen KIs bevorzugt werden.

Bottom-Up Prinzip in der Entwicklung

Grundsätzlich soll die Entwicklung der KI einem Bottom-Up-Prinzip folgen.

Einsatz KI je nach Anwendungsbereich

Auch hier wird für einen Einsatz der KI als Datenbank je nach Anwendungsbereich appelliert, wobei Chancen vor allem im pathologischen Bereich identifiziert werden. Eine Vorselektierung der Daten erscheint hier sinnvoll.

Einfachheit der Nutzung

Sicherstellung und Nachvollziehbarkeit der Datensicherheit

Kompatibilität

Standardisierung

Berücksichtigung rechtlicher Vorgaben für die (emotionale) Gestaltung von KI

Weitere Anforderungen sind die Einfachheit der Nutzung, die Sicherstellung und Nachvollziehbarkeit der Datensicherheit, die Kompatibilität und

Standardisierung der Daten, sowie die Berücksichtigung rechtlicher Vorgaben für die (emotionale) Gestaltung von KI.

Digitalkompetenz: Einfachheit und Verständlichkeit

Ein weiterer Punkt ist die auf die Digitalkompetenz der Nutzenden ausgerichtet. Die KI sollte einfach und verständlich sein. Dabei soll der Nutzungskontext eine Unterstützung zum Erwerb digitaler Kompetenz liefern können, sowie Barrierefreiheit beispielsweise in Form von einfacher Sprache garantiert werden.

Vertrauensbildende KI Entwicklung

Einen größeren Themenblock stellt die vertrauensbildende KI Entwicklung dar. Hier wurden folgende Anforderungen genannt: das Einführen eines vertrauenswürdigen Datenökosystems in der EU, das Einhalten der European Digital Rights and Principles, öffentliche Nachvollziehbarkeit und Vermeidung von Blackbox, das Einführung eines Ampelsystems, Fokus auf Selbstbestimmung und Agency, Vermeidung von Manipulation (v.a. durch Pharmakonzerne), sowie die Entwicklung einer verantwortungsvollen, gemeinwohlorientierten Technologie.

Frage nach Emotionalisierung der KI

Ratio über Emotio basierte Entscheidungsfindung

Eine Rolle spielt weiterhin die Frage nach der Emotionalisierung der KI. Grundsätzlich wurde jedoch festgestellt, dass die Entscheidungsfindung der KI eher Ratio als Emotio basiert sein soll. Der Mensch muss dabei im Hintergrund bestehen bleiben.

Berücksichtigung des Spaßfaktors der Mensch-Technik-Interaktion

Bei der Entwicklung sollte der Spaßfaktor der Mensch-Technik-Interaktion berücksichtigt werden.

Blackbox: Backtracking Funktion zur späteren Überprüfbarkeit

Um eine Blackbox zu vermeiden, muss es eine Backtraining Funktion zur späteren Überprüfbarkeit der KI im Sinne einer explainable AI geben.

Einführung einer Kennzeichnung (Digital-Responsibility-Index)

Abschließend sollte eine Kennzeichnung, basierend auf dem Digital-Responsibility-Index und ähnlich der Energieeffizienzklassifizierung von

Elektrogeräten, eingeführt werden. Folgende Leitlinien sind zu berücksichtigen: Cybersicherheit, Selbstbestimmung, digitale Kompetenz, Transparenz.

Im Folgenden sind die Anforderungen zum Akzeptanzobjekt und spezifisch für Ärzte zutreffend dargestellt:

KI als Unterstützung für Ärzte bei Entscheidungsfindung

Hier genannt wurde die Anforderung, KI als Unterstützung bei der Entscheidungsfindung zu entwerfen. Dabei soll die KI die Fehlerquote reduzieren. Es wurde die Frage der Gewichtung von KI und Mensch im Entscheidungsprozess aufgeworfen, sowie, ob der Einbezug einer KI zur interdisziplinären Entscheidungsfindung legitim sei.

KI als Arbeitserleichterung für Ärzte bei großer Datenmenge

Des Weiteren soll die KI eine Arbeitserleichterung für Ärzte bei großen Datenmengen darstellen.

Einbezug von Ärzten bei der Entwicklung und Wissensgenerierung

Ärzte sollen bei der Entwicklung und Wissensgenerierung der KI einbezogen werden.

3.2.4 Implementierung: Akzeptanzkontext und Akzeptanzsubjekt

Einen letzten thematischen Punkt der Gruppendiskussion stellt die Implementierung der KI dar. Hierunter fallen alle genannten Inhalte und Anforderungen, welche für eine letztendliche Implementierung der Technologie von Bedeutung sind. Auch hier fand eine Unterteilung in Anlehnung an Lucke (1995) statt. Inhalte des Themenblocks decken sich mit dem oben genannten Begriff Individuum und Nutzungskontext, beziehungsweise Akzeptanzsubjekt und Akzeptanzkontext. Auch hier dienen die Begriffe als Übercodes. Insgesamt wurden 43 Codes zur Implementierung, darunter 37 zum Akzeptanzkontext, gezählt. Auch hier fand die Unterteilung in „Allgemein“ mit 26 Codes und „spezifisch auf Ärzte zutreffend“ mit 11 Codes statt. Die zweite thematische Ausrichtung stellen Inhalte bezogen auf das Akzeptanzsubjekt mit insgesamt 6 Codes, darunter 4 für „Allgemein“ und 2 für spezifisch auf Ärzte zutreffend“, dar.

3.2.4.1 Akzeptanzkontext

Im Folgenden sind die allgemein zutreffenden Anforderungen zum Akzeptanzkontext dargestellt:

(Technische) Transparenz auf Nachfrage

Eine wichtige Anforderung stellt hier die (technische) Transparenz auf Nachfrage dar. Darunter zählen die Zielgruppenabhängige Transparenz, Transparenz in Bezug auf den Gesamt-Kontext und nicht ausschließlich auf KI Inhalte. Eine Anforderung ist die gleiche Möglichkeit der Transparenz für alle Personengruppen, dabei jedoch unterschiedliche Ebenen der Transparenz, vor allem bei der Unterscheidung zwischen Ärzten und Patienten. Transparenz ist für Ärzte von größerer Bedeutung, als für Patienten beziehungsweise Versicherte. Hinzu kommt die (technische) Nachvollziehbarkeit auch für Laien, Transparenz über die Datenverarbeitung und Datengrundlage. Zum Thema Transparenz wurde die Einschränkung getroffen, dass ohne entsprechendes technisches Wissen KI nicht vollständig nachvollziehbar sei. Vollkommene Transparenz könnte jedoch zu Verwirrung führen, weshalb ein unterschiedliches Ausmaß an Transparenz für unterschiedliche Personengruppen als sinnvoll erachtet wird.

Berücksichtigung: darf eine KI nachtrainiert werden?

Ebenfalls wichtig ist, dass berücksichtigt wird ob eine KI nachtrainiert werden darf.

Berücksichtigung der Rolle von Medien bezüglich KI Akzeptanz

Kommunikation des Gewinns und Nutzens der KI

Kommunikation von Informationen über KI in richtigem Ausmaß

Durch Kommunikation und Transparenz Vertrauen schaffen

Weitere Anforderungen sind die Berücksichtigung der Rolle von Medien bezüglich der Akzeptanz von KI, dabei auch die Kommunikation des Gewinns und Nutzens der KI. Informationen über die KI sollen jedoch im richtigen Ausmaß und letztlich nutzungssituationsabhängig kommuniziert werden. Durch Kommunikation und Transparenz soll letztlich Vertrauen geschaffen werden.

Versicherte haben letzte Kontrollinstanz über Daten

Versicherte haben Wahlfreiheit über Nutzung von KI in Krankenkassen

Weiterhin wichtig ist, dass Versicherte die letzte Kontrollinstanz über die eigenen Daten haben und über die Nutzung von KI vor allem in Krankenkassen frei wählen dürfen.

Im Folgenden sind die spezifisch für Ärzte zutreffenden Anforderungen zum Akzeptanzkontext dargestellt:

Nutzerzentrierte Implementierung

Anforderungen sind die nutzerzentrierte Implementierung unter Berücksichtigung der Haftungsfrage als Akzeptanzbarriere. Beispiele sind Fort- und Weiterbildungen, Aufklärungsarbeit, sowie der Einbezug von Ärzten.

Kein zwanghaftes Durchsetzen der Implementierung

Grundsätzlich sollte die Implementierung der KI nicht zwanghaft durchgesetzt werden, sondern einem Top-Down Prinzip folgen. Zudem sollte eine Feedback-Funktion eingeführt werden.

Mitbestimmung von Ärzten bei der Implementierung

Wie bereits angemerkt, ist die Mitbestimmung von Ärzten bei der Implementierung von großer Bedeutung. Als Beispiel wurden Demonstratoren für Kongresse und ähnliches angeführt.

Kommunikation von Nutzen der KI

Auch hier wird die Kommunikation des Nutzens der KI angeführt.

Positiver Einfluss von KI Systemen auf Lernen und Wissensproduktion

Letztlich wurde in der Diskussion angeführt, dass KI Systeme einen positiven Einfluss auf Lernen und die Wissensproduktion nehmen können. Aus diesem Grund soll die KI mit einer Lern- und Lehrfunktion ausgestattet sein. Zudem soll so die Gefahr von Verunsicherung des Anwenders durch beispielsweise Vorwissen, Intuition und Erfahrung vermieden werden.

3.2.4.2 Bezogen auf Akzeptanzsubjekt

Im Folgenden sind die allgemein zutreffenden Anforderungen bezogen auf das Akzeptanzsubjekt dargestellt:

Berücksichtigung von Vorerfahrung/ Vorkenntnissen mit KI

Unterscheidung: Versicherte sind nicht gleich Patienten

Anforderungen in diesem Themenbereich sind die Berücksichtigung von Vorerfahrung und Vorkenntnissen von Nutzenden mit der KI. Zudem soll eine grundlegende Unterscheidung zwischen Patienten und Versicherten getroffen werden.

Höhere Akzeptanz Falschaussage Mensch als richtige Aussage von Maschine

Grundsätzlich wird eine höhere Akzeptanz der Falschaussage eines Menschen, als einer richtigen Aussage der KI angenommen.

Fokussierung auf Vertrauen und Transparenz

Zuletzt soll sich eher auf KI Vertrauen, als auf Transparenz fokussiert werden. Erwähnt wurde, dass Vertrauen vor allem auf Versicherten-Ebene insgesamt einen höheren Stellenwert hat als Transparenz. Somit soll das Ziel verfolgt werden eine "vertrauensvolle Instanz", die transparent, ethisch und regulatorisch ist, zu schaffen.

Im Folgenden sind die spezifisch auf Ärzte zutreffenden Anforderungen bezogen auf das Akzeptanzsubjekt dargestellt:

Berücksichtigung von Interaktionseffekten für Akzeptanz

Die Anforderung hierfür ist die Berücksichtigung von Interaktionseffekten für die Akzeptanz. Im Zuge dessen wurde die Frage aufgeworfen, ob beispielsweise das Alter und Maß der Informiertheit eine Rolle spielen könnten. Dies wurde im Laufe der Diskussion jedoch verworfen.

3.3 Ergebnisse der Priorisierung

Im Anschluss an die Diskussion der Anforderungen in Kleingruppen, sollten die Teilnehmer in einer letzten gemeinsamen Gruppendiskussion zentrale Aspekte auf Post-Its festhalten (siehe 7.1). Diese können als Priorisierung gesehen werden, da es sich um einschlägige, am wichtigsten empfundene Aspekte handelt. Folgende 17 Stichpunkte wurden auf Post-Ist festgehalten:

- Human Agency first
- Europäisch denken + handeln
- Digital & Health Literacy: Transparenz, einfache Sprache, inklusiv + intuitiv!
- KI und ÄrztInnen sollten sich gegenseitig "auf Augenhöhe" ergänzen, hinterfragen, challengen
- KI darf keine Black-Box sein: Erklärungskomponente, Nachvollziehbarkeit
- KI ist nicht KI. Um welche Methoden geht es konkret?
- Zentrale Dimensionen bei KI-Akzeptanz sollten nicht nur Nützlichkeit, Einfachheit oder Intention berücksichtigen, sondern vor allen Dingen Zuverlässigkeit/Verbindlichkeit
- Verantwortung und Entscheidungshoheit bleibt in ärztlicher Hand
- Usability
- Implementierung/Infrastruktur
- Nutzer:innen Perspektive berücksichtigen
- Aufklärung der Zielgruppen: Kompetenz + Verständnis erzeugen
- Zielgruppen Befragung zu "Was wird gebraucht?": Gap Analyse, Zielgruppe mitnehmen in der KI-Entwicklung
- Große Dynamik, viele Baustellen/ungeklärte Fragen (Recht, Ethik, Gesellschaft)
- Zentrale Stelle von Kontrolle + Vertrauen (bei UI + Medizin)
- Möglichst einheitliches deutsches/europäisches System, um Datenschutz + Kompatibilität zu garantieren
- Gute Strategien zur Implementierung nötig, sonst hilft alle Akzeptanz nichts: "Mitspracherecht" von ÄrztInnen & Versicherten

Hier spiegeln sich zum Teil oben genannte Anforderungen und Bedenken wieder. Zudem lassen sich die genannten Aspekte ebenfalls in die vier identifizierten Themenbereiche einordnen, sodass sich aus den schriftlich festgehaltenen Stichpunkten folgende Ableitung ergibt:

Grundlegende Definition von KI:

- Human-Agency first
- Begegnung von KI und Ärzt:innen "auf Augenhöhe" (ergänzen, hinterfragen, challengen)
- Entscheidungshoheit liegt bei Ärzt:innen

Grundlegende Rahmenbedingungen von KI:

- einheitliches System
- deutsche/europäische Orientierung bezüglich Datenschutz und Komptabilität (zweifache Nennung)

Entwicklung:

- Transparenz
- einfache Sprache
- inklusiv
- intuitiv und einfach
- nachvollziehbar (Erklärungskomponente)
- zuverlässig/verbindlich
- Usability
- Berücksichtigung der Nutzer:innen Perspektive

Implementierung:

- Gute (Infra-)Struktur und Strategie (zweifache Nennung)
- Einbindung der Zielgruppen (Befragung, Aufklärung) (dreifache Nennung)

Zudem konnten folgende Aspekte als Bedenken, welche für die Entwicklung von Bedeutung identifiziert werden:

- Um welche Methoden im Bereich KI geht es konkret?
- Große Dynamik: viele Baustellen und ungeklärte Fragen (Recht, Ethik, Gesellschaft)
- Kontrolle und Vertrauen als zentraler Faktor.

Im Folgenden werden die Ergebnisse abschließend zusammengeführt und reflektiert.

4 Zusammenführung

Anhand der obigen Auswertungen lässt sich feststellen, dass Anforderungen in unterschiedlichen Bereichen identifiziert werden konnten. Dabei wurde deutlich, dass einige Anforderungen direkt in die Praxis ableitbar sind, wohingegen es sich bei anderen um eher grundlegende Aspekte handelt. Zudem lässt sich feststellen, dass manche Themen insgesamt wichtiger zu sein scheinen, als andere. Dies lässt sich zum einen aus der Häufigkeit der Erwähnungen und der Ähnlichkeit mancher Aspekte ableiten, zum anderen aus der Priorisierung der Teilnehmenden, welche am Ende des Workshops auf Post-ist festgehalten wurden (siehe 3.3).

Zu den eher **grundlegenden Überlegungen und Aspekten** zählen vor allem Anforderungen aus dem Themenbereich „**grundlegende Definition von KI**“ und „**grundlegende Rahmenbedingungen**“ (siehe 3.2.1 und 3.2.2). Dabei ist klar, dass es eine eindeutige Definition von KI, der Rolle des Menschen in Auseinandersetzung mit der KI und die Festlegung der Gewichtung der KI innerhalb des medizinischen Kontexts braucht. Zudem müssen Rahmenbedingungen geklärt werden, vor allem da die Anwendung von KI im pathologischen Bereich besser als beispielweise im therapeutische Bereich bewertet wurde. Hierauf bezogen scheint das Thema Kontrolle und Vertrauen von Bedeutung zu sein. Auch in der Priorisierung taucht die „zentrale Rolle von Kontrolle und Vertrauen in KI und Medizin“ auf. Zudem scheint es von Bedeutung zu sein, dass KI und Ärzte sich „auf Augenhöhe“ begegnen.

In diesem Themenblock handelt es sich um noch offene und zu klärende Fragen. In der Kombination mit der Priorisierung lässt sich zusammenfassend ableiten, dass bei der Einführung von KI im medizinischen Bereich große Dynamik mit vielen Baustellen und ungeklärten Fragen (Recht, Ethik, Gesellschaft) herrscht. Die letzte Entscheidungsinstanz muss jedoch, hier waren sich alle Teilnehmende einig, letztendlich bei Ärzten bzw. dem medizinischem Personal und nicht bei der KI liegen.

Weitere offene Fragen beziehen sich auf rechtliche und wirtschaftliche Faktoren. Darunter zählt der mehrfach genannte und bei der Priorisierung festgehaltene Wunsch der Standardisierung der KI innerhalb der EU, sowie deren wirtschaftliche und politische Unabhängigkeit und Neutralität. Besonders von Bedeutung scheint die Standardisierung innerhalb der EU, sowie das Einhalten europäischer Richtlinien zu sein, da die Punkte sowohl in der Diskussion mehrfach auftauchten, als auch in der Priorisierung: „europäisches denken und handeln“, „europäisches System“.

Direkt in die Praxis umsetzbare Anforderungen beziehen sich auf die Entwicklung (Akzeptanzobjekt) und Implementierung (Akzeptanzsubjekt und -kontext) (siehe 3.2.3 und 3.2.4). In diesem Kontext besonders hervorzuheben sind die nachfolgenden Aspekte: Interessant ist zunächst, dass explizit auf die Digital Responsibility Goals (DRGs) eingegangen wird. Diese gilt es im Zuge der Entwicklung zu erreichen. Neben der expliziten Nennung der DRGs decken sich auch weitere diskutierte Anforderungen mit diesen. Hierunter zählt vor allem der

detailliert ausgearbeitete Aspekt der Transparenz, sowie die Nennung von vertrauenswürdigen Algorithmen und digitaler Kompetenz. Die Einhaltung der Digital Responsibility Goals bei der Entwicklung von KI scheint für die Teilnehmenden von Bedeutung zu sein. So tauchen eben genannte Aspekte auch in der Priorisierung auf (Transparenz, Aufklärung von Zielgruppen).

Ein wichtiges Thema und extra hervorzuheben ist zudem der Umgang mit (technischer) Transparenz. Das Thema wurde an verschiedener Stelle (3.2.1, 3.2.3, 3.2.3) genannt und taucht auch in der Priorisierung der Teilnehmenden auf. Transparenz ist somit für die Definition von KI, das Akzeptanzobjekt selbst, als auch dessen Akzeptanzkontext, sowie bezogen auf das Akzeptanzsubjekt, von Bedeutung. Aus der Diskussion und der Priorisierung lässt sich ableiten, dass Transparenz zielgruppenabhängig sein sollte, jedoch grundsätzlich jedem zugänglich und nachvollziehbar sein muss. Das deckt sich auch mit in der Priorisierung, in welcher die Anmerkung „Digital und Health literacy: Transparenz, einfache Sprache, Inklusion und intuitiv“ auftaucht.

Ein weiteres Thema, welches im Zuge der Auswertung häufiger auftaucht ist die Kommunikation über KI. Dies bezieht sich ausschließlich auf den Akzeptanzkontext und ist somit für die Implementierung von Bedeutung (siehe 3.2.4.1). Festzustellen ist, dass Kommunikation über KI im richtigen Ausmaß und nutzungssituationsabhängig stattfinden sollte. Kommunikation ist letztlich mit dem Thema Transparenz verschränkt, da es sich bei richtiger Kommunikation laut den Teilnehmenden primär um Aufklärungsarbeit handelt.

Im Zuge der Auswertung konnte abschließend festgestellt werden, dass es bei den Teilnehmern bereits eine Vorstellung davon gibt, was KI im medizinischen Kontext sein sollte. Dies basiert auf den unter 3.2.1 und 3.2.3 festgestellten Punkten, ist also auf die Definition von KI und das Akzeptanzobjekt bezogen. Die folgenden Faktoren könnten eine erste Antwort auf die Forderung nach einer Definition von KI darstellen. Zusammenfassend sollte die KI im medizinischen Kontext folgendes sein:

KI als...

- kritische Hinterfragungskomponente
- Algorithmus
- Empfehlung im Entscheidungsprozess
- Statistikpool
- Werkzeug
- Datenbank je nach Anwendungsbereich
- Unterstützung der Entscheidungsfindung
- KI im Sinne einer explainable AI geben

KI zur...

- Fehlerreduktion
- Reflexion von Arbeitsschritten
- Einsatz von KI in der Lehre und zur Schulung von medizinischem Personal
- Verringerung der Arbeitsbelastung für Ärzte

Anhang 5: Ergebnisdokumentation AP 9 – Handlungsempfehlungen

- Steigerung der Zeiteffektivität
- Verbesserung der Prozessstrukturen
- Steigerung der diagnostischen Sicherheit
- Arbeitserleichterung bei großen Datenmengen

5 Fazit

Zusammenfassend lässt sich festhalten, dass nicht nur Handlungsempfehlungen, sondern konkrete Anforderungen an eine erfolgreiche Entwicklung und Implementierung einer KI identifiziert werden konnten.

Die Aspekte Definition von KI, Definition der Rahmenbedingungen, Entwicklung und Implementierung (siehe Abbildung 1) können als Anhaltspunkt betrachtet werden. Zunächst muss eine Grundlegende Definition von KI, sowie die Definition der Rahmenbedingungen für die Entwicklung und Implementierung von KI entworfen werden. Die Teilnehmer haben bereits erste Vorstellungen zu einer Definition von KI. Hierfür ist vor allem der Arbeitsbereich Medizin, sowie die Rolle des Menschen in Bezug auf die KI von Bedeutung. Bei letzterem spielen rechtliche und wirtschaftliche Bedingungen eine primäre Rolle.

Anschließend können Anforderungen an die Entwicklung und Implementierung von KI entworfen werden. Während sich ersteres auf Anforderungen bezüglich des Akzeptanzobjekts bezieht, sind bei letzterem Akzeptanzkontext und Akzeptanzsubjekt von Bedeutung.

Abschließend muss angemerkt werden, dass bei der Ausrichtung von Anforderungen auf Handlungsempfehlungen zwei Zielgruppen unterschieden werden. Dabei handelt es sich um allgemein zutreffende und spezifisch auf Ärzte zutreffende Anforderungen.

Die vier identifizierten Themenbereiche können nicht gänzlich getrennt voneinander gesehen werden. Sie sind zum Teil miteinander verschränkt. Die Berücksichtigung dieser Punkte wirkt sich letztlich positiv auf die Akzeptanz einer KI aus. Weiterführend könnte betrachtet werden, inwiefern sich die identifizierten Aspekte zusätzlich auf Vertrauen und Kontrolle auswirken.

6 Literatur

Hackel, M., & Klebl, M. (2008). Qualitative Methodentriangulation bei der arbeitswissenschaftlichen Exploration von Tätigkeitssystemen. *Forum Qualitative Sozialforschung*, 9(3).

Lucke, D. (1995): Akzeptanz. Legitimität in der "Abstimmungsgesellschaft". Opladen: Leske + Budrich.

Mayring, P., & Fenzl, T. (2019). Qualitative Inhaltsanalyse. In: Baur, N., Blasius, J. (eds) *Handbuch Methoden der empirischen Sozialforschung*. Springer VS, Wiesbaden.

Neyer, F. J., Felber, J., & Gebhardt, C. (2016). Kurzsкала Technikbereitschaft (TB, technology commitment). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS).

Tausch, A., & Menold, N. (2015). Methodische Aspekte der Durchführung von Fokusgruppen in der Gesundheitsforschung. 51.

7 Anhang

7.1 Workshop-Dokumentation

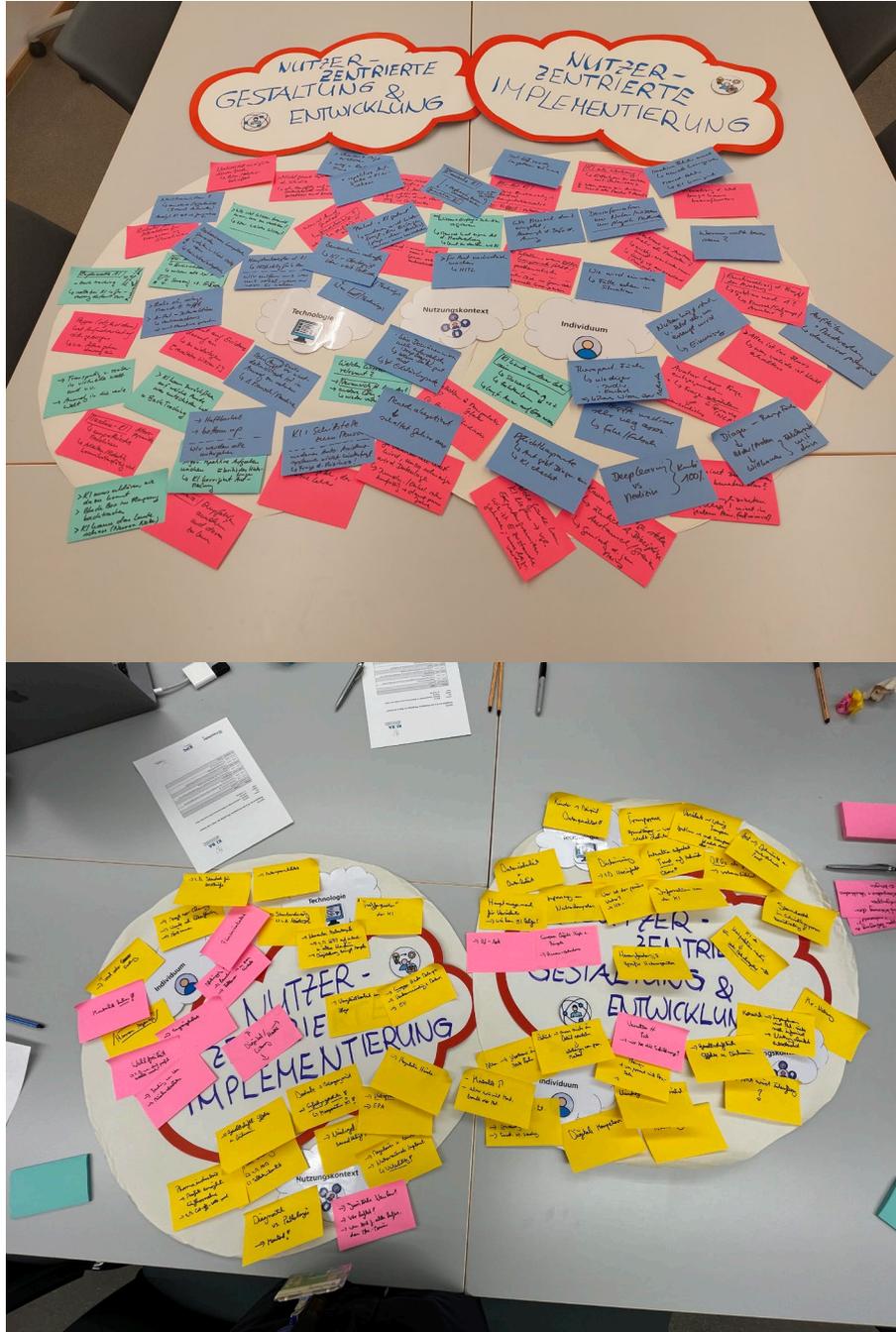


Abbildung 1: Ergebnisse aus den Gruppendiskussionen

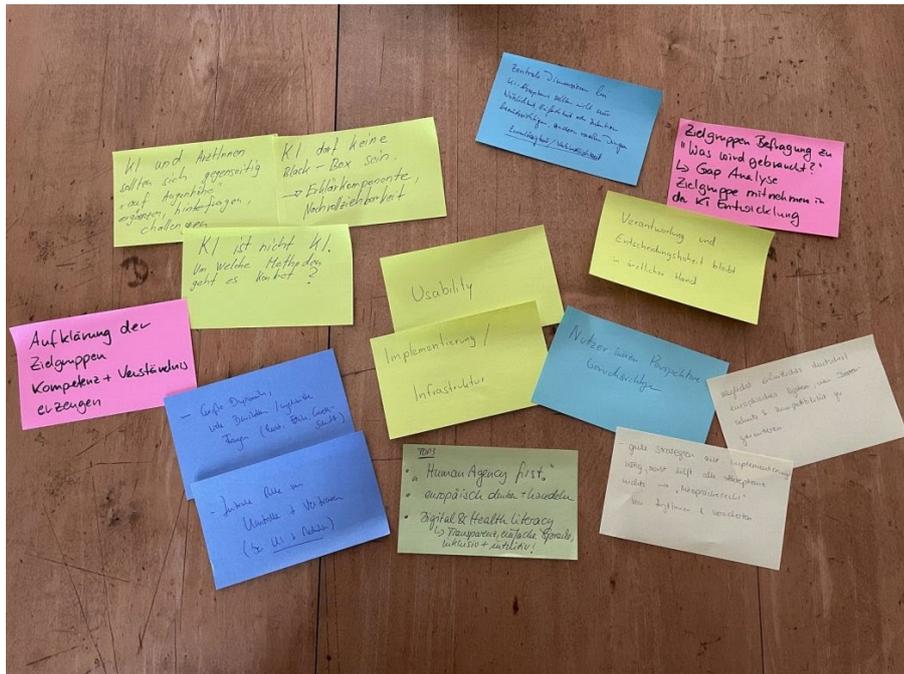


Abbildung 2: Priorisierung

7.2 MAXQDA-Codesystem

<ul style="list-style-type: none"> ▼ Codesystem ▼ Handlungsempfehlung ▼ Grundlegende Definition von KI > Allgemein > spezifisch für Ärzte zutreffend ▼ Grundlegende Rahmenbedingungen > Allgemein > spezifisch für Ärzte zutreffend ▼ Entwicklung ▼ Akzeptanzobjekt > Allgemein > spezifisch für Ärzte zutreffend ▼ Implementierung ▼ Akzeptanzkontext > Allgemein > spezifisch für Ärzte zutreffend ▼ Bezogen auf Akzeptanzsubjekt > Allgemein > spezifisch für Ärzte zutreffend 	<p>142</p> <p>0</p> <p>0</p> <p>15</p> <p>19</p> <p>0</p> <p>10</p> <p>6</p> <p>0</p> <p>0</p> <p>39</p> <p>10</p> <p>0</p> <p>0</p> <p>26</p> <p>11</p> <p>0</p> <p>4</p> <p>2</p>
---	---

Abbildung 3: Übersicht Codesystem

Anhang 5: Ergebnisdokumentation AP 9 – Handlungsempfehlungen

Entwicklung	0
Akzeptanzobjekt	0
Allgemein	0
Generierung der Datenbank der KI als wichtiger Punkt	1
Problematik: Aktualität (Veraltetes Wissen und Daten)	1
Woher kommen Daten (Validität, Güte?)	1
Ausgereifte, valide Datenbank für Implementierung	3
Zentrale Bedeutung valider Daten	1
Beförderung analytischer statt generativer KI	1
Bottom-Up Prinzip in der Entwicklung	1
Einsatz KI je nach Anwendungsbereich	1
Vorteile/Möglichkeit vor allem in der Pathologie gegeben	1
Einfachheit der Nutzung	1
Sicherstellung und Nachvollziehbarkeit der Datensicherheit	2
Kompatibilität	1
Standardisierung	1
Berücks. rechtlicher Vorgaben f. (emotionale) Gestaltung v. KI	1
Digitalkompetenz: Einfachheit und Verständlichkeit	1
Nutzungskontext als Unterstützung bei digitaler Kompetenz	1
Barrierefreiheit: Einfache Sprache	1
vertrauensbildende KI Entwicklung	1
vertrauenswürdige Datenökosystem in EU	1
Einhalten der European Digital Rights and Principles	1
Öffentliche Nachvollziehbarkeit, Vermeidung von Blackbox	1
Einführung eines Ampelsystems	1
Fokus auf Selbstbestimmung und Agency	1
Vermeidung von Manipulation (v.a. durch Pharmakonzerne)	1
responsible Technology, gemeinwohlorientiert	1
Frage nach Emotionalisierung von KI	1
Ratio über Emotio basierte Entscheidungsfindung	3
Berücksichtigung des Spaßfaktors d. Mensch-Technik-Interaktion	1
Blackbox: Backtracking Funktion zur späteren Überprüfbarkeit	1
Einführung einer Kennzeichnung (Digital-Responsibility-Index)	1
Leitlinien: Cybersicherheit	1
Leitlinien: Selbstbestimmung	1
Leitlinien: digitale Kompetenz	1
Leitlinien: Transparenz	1
spezifisch für Ärzte zutreffend	0
Einbezug v. Ärzten bei d. Entwicklung u. Wissensgenerierung	1
KI als Arbeitserleichterung f. Ärzte bei großer Datenmenge	1
KI als Unterstützung für Ärzte bei Entscheidungsfindungen	3
KI als Unterstützung bei der Entscheidungsfindung	2
Frage nach Gewichtung KI und Mensch in Entscheidungsprozess	1
KI zur Reduzierung der Fehlerquote	2

Abbildung 6: Codes Entwicklung

Implementierung	0
Akzeptanzkontext	0
Allgemein	0
(technische) Transparenz auf Nachfrage	7
Zielgruppenabhängige Transparenz	1
Transparenz in Bezug auf Gesamt-Kontext nicht KI-Inhalte	2
gleiche Möglichkeit der Transparenz für alle Personengruppen	1
Unterschiedliche Ebenen der Transparenz: Ärzte/Patienten	1
Transparenz für Ärzte von großer Bedeutung	1
Transparenz für Patienten/ Versicherte weniger bedeutend	2
(technische) Nachvollziehbarkeit auch für Laien	1
Transparenz über Datenverarbeitung	1
Transparenz über Datengrundlage	1
Berücksichtigung: darf eine KI nachtrainiert werden?	1
Berücksichtigung der Rolle von Medien bezüglich KI Akzeptanz	1
Kommunikation des Gewinns und Nutzens der KI	1
Kommunikation von Informationen über KI in richtigem Ausmaß	1
Nutzungssituationsabhängig	1
Durch Kommunikation und Transparenz Vertrauen schaffen	1
Versicherte haben letzte Kontrollinstanz ü. Daten	1
Versicherte haben Wahlfreiheit ü. Nutzung v.KI in Krankenkassen	1
spezifisch für Ärzte zutreffend	0
Nutzerzentrierte Implementierung	1
Berücksichtigung: Haftung als Akzeptanzbarriere	1
Kein zwanghaftes Durchsetzen der Implementierung	2
Top-Down Prinzip für die Implementierung	1
Einführen einer Feedback-Funktion	1
Mitbestimmung von Ärzten bei der Implementierung	1
Kommunikation von Nutzen der KI	1
pos. Einfluss von KI Systemen auf Lernen und Wissensproduktion	1
KI mit Lern- und Lehrfunktion	1
Vermeidung der Gefahr von Verunsicherung des Anwenders	1
Bezogen auf Akzeptanzsubjekt	0
Allgemein	0
Berücksichtigung von Vorerfahrung/Vorkenntnissen mit KI	1
Unterscheidung: Versicherte sind nicht gleich Patienten	1
höhere Akzeptanz Falschaussage Mensch als richtige v. Maschine	1
Fokussierung auf Vertrauen und weniger auf Transparenz	1
spezifisch für Ärzte zutreffend	0
Berücksichtigung von Interaktionseffekten f. Akzeptanz	1
Einfluss von Alter und Informiertheit nicht vorhanden	1

Abbildung 7: Codes Akzeptanzkontext

Abbildungsverzeichnis

Abbildung 2: Ergebnisse aus den Gruppendiskussionen	26
Abbildung 3: Priorisierung.....	27
Abbildung 4: Übersicht Codesystem	27
Abbildung 5: Codes Grundlegende Definition von KI.....	28
Abbildung 6: Codes Grundlegende Rahmenbedingungen.....	28
Abbildung 7: Codes Entwicklung.....	29
Abbildung 8: Codes Akzeptanzkontext.....	29

Tabellenverzeichnis

Tabelle 1: Ergebnisse Kurzsкала Technikbereitschaft

8

KI-BA:

Künstliche Intelligenz in der Versorgung – Bedingungen der Akzeptanz von Versicherten und ÄrztInnen

**Manual Hauptstudie – Ärzte Stand:
21.10.2022**

Projektmitarbeitende:

Fraunhofer IIS: M. Simon, S. Schmitt-Rüth, H. Japp, B. Williger, & T. Wittenberg IPG: S. T. Kamin, H. S. Hösch, N. M. Degen, A. Beyer & F. R. Lang

FORSCHUNGSDESIGN	4
EINLEITUNGSTEXT	5
DATENSCHUTZ	7
IDCODE	8
SOZIODEMOGRAPHISCHE FAKTOREN	9
ALTER UND GESCHLECHT	9
RÄUMLICHER KONTEXT	10
GEMEINDEGRÖÖE	10
POSTLEITZAHL	10
BERUF	11
FACHARZT & AUSBILDUNG.....	11
BERUFLICHE POSITION	12
BESCHÄFTIGUNG.....	13
ARBEITSBEDINGUNGEN.....	14
ARBEITSBELASTUNG.....	14
ZEIT FÜR PATIENTINNEN	15
ARZT-PATIENTEN-BEZIEHUNG.....	15
BERUFSWECHSEL.....	16
DIGITALISIERUNG	16
SCHULUNGEN.....	17
AUSSTATTUNG	17
KI-ERFAHRUNGEN.....	18
TECHNIKMOTIVATION	19
VIDEO – EINFÜHRUNG	21
KI-AKZEPTANZ IN DER GESUNDHEITSVERSORGUNG.....	22
ALLGEMEINE KI-AKZEPTANZ.....	24
VIDEO – STIMULI	25
MANIPULATIONSCHECK	26
TECHNIKAKZEPTANZ.....	28
PERSÖNLICHKEIT	29
PERSÖNLICHKEITSDISPOSITIONEN	29
ZEITPERSPEKTIVE.....	31
OBSOLESZENZERLEBEN	32
SUBJEKTIVES ÄLTERSERLEBEN	33
NOVELTY.....	34
NEUGIER.....	35
LEBENSZUFRIEDENHEIT	36
GESUNDHEITLICHE RESSOURCEN	37
SUBJEKTIVE GESUNDHEIT.....	37
EINSAMKEIT.....	38
ÄNGSTLICHKEIT UND DEPRESSION	39
GESUNDHEITSVERHALTEN	40
VORSORGE	40
SOZIALE BEZIEHUNGEN	41
ZUFRIEDENHEIT MIT BEZIEHUNGEN	41
SOZIALE UNTERSTÜTZUNG	42
ANZAHL WICHTIGER BEZIEHUNGEN	43
ABSCHLUSS DER BEFRAGUNG	44
RELUKTANZ	44
ZUVERLÄSSIGKEIT	45
ABLENKUNG	46
ANMERKUNGEN.....	47
KONTAKT	47

Forschungsdesign

Kurzbeschreibung:

Grundanliegen des Vorhabens ist es zu verstehen, wie die Akzeptanz gegenüber KI-Systemen bei Versicherten und ÄrztInnen ausgeprägt ist und welche Einflussfaktoren die individuelle Akzeptanz beeinflussen. Dafür werden den Probandinnen und Probanden im Rahmen eines 2x2-faktoriellen Designs randomisierte KI-Anwendungsszenarien in Form animierter Kurzvideos präsentiert. Die Videos repräsentieren die zuvor entwickelte Kategorisierung und sind mit Blick auf die individuelle Akzeptanz gegenüber der entsprechenden KI-Anwendung zu bewerten.

Versuchsplan:

Der mehrfaktorielle Versuchsplan beinhaltet zwei unabhängige Variablen mit jeweils zwei Faktorstufen und eine unabhängige Variable mit 3 Faktorstufen (2x2x3 Design):

- IV1: AI-Explainability (high/low), d.h. die Funktionsweise der KI wird dargestellt – oder nicht
- IV2: AI-Completeness (high/low), d.h. die KI entscheidet autonom – oder der Mensch
- IV3: KI Use-case (Diabetes, EKG, Koloskopie)

Aufbau der Befragung:

Der Aufbau der Befragung ist in Abb. 1 dargestellt. Die erste Seite umfasst die Erhebung grundlegender soziodemographischer Informationen. Danach werden die Probanden gebeten, sich ein Video zur Einführung in die KI-Thematik anzuschauen. Nach diesem Video erfolgt die Erhebung der KI- Akzeptanzskala, die im Rahmen des Projektes entwickelt wurde, um interindividuelle Unterschiede in der Bewertung der KI-Lösungen zu erklären. Danach erfolgt die randomisierte Zuteilung zur einer between- subject Bedingung, innerhalb derer eine randomisierte Videosequenz zu evaluieren ist. Die Evaluation erfolgt anhand klassischer Instrumente aus der Technikakzeptanz- sowie der UX-Forschung, die gleichzeitig als zentrale abhängige Variablen fungieren. Nach der Evaluationen werden die Probanden gebeten, weitere Fragen zu potentiellen Einflussfaktoren der KI-Akzeptanz zu beantworten. Dabei handelt es sich um die Bereiche:

- Räumlicher Kontext (z.B. Versorgungssituation, Gemeindegröße)
- Soziale Beziehungen (z.B. Anzahl Beziehungspartner, soziale Unterstützung)
- Gesundheitliche Ressourcen (z.B. Krankheiten, Alltagsfunktionalität)
- Gesundheitsverhalten (z.B. Vorsorgeverhalten, Arzt-Patienten-Beziehung)
- Persönlichkeit (z.B. Zeitperspektive, Alterserleben)
- Technik und Digitalisierung (z.B. Technikmotivation, Digitalkompetenz)

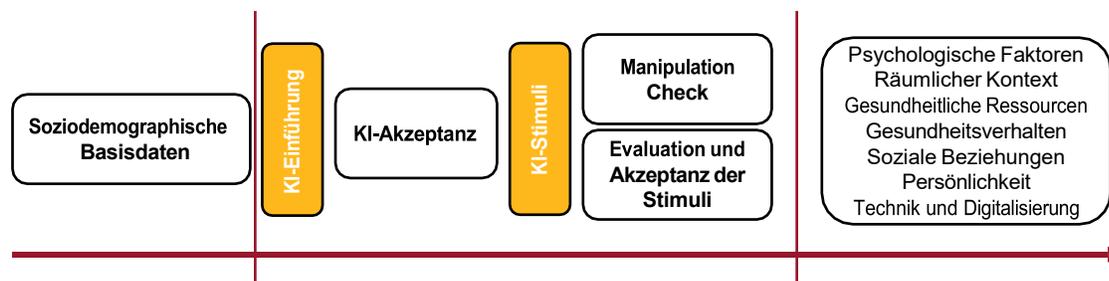


Abbildung 1: Aufbau der Befragung

Im Folgenden sind Variablennamen rot und Variablenwerte grün gekennzeichnet. Textblöcke aus dem Fragebogen sind grau hinterlegt.

Einleitungstext

Herzlich willkommen zur Online-Befragung des Projekts "KI-BA"
"Künstliche Intelligenz (KI) in der Versorgung - Bedingungen der Akzeptanz von Versicherten und ÄrztInnen"

Was ist das Ziel unserer Studie?

Mit Hilfe künstlicher Intelligenz können in der Gesundheitsversorgung besonders komplexe Abläufe und Entscheidungsprozesse optimiert werden. Unklar ist allerdings bisher, in welcher Form diese Technologie im Gesundheitsbereich Anwendung finden kann. Viele Faktoren beeinflussen die Umsetzung und Nutzung künstlicher Intelligenz, nicht zuletzt spielt die Akzeptanz aller Beteiligten eine wesentliche Rolle.

Mit Hilfe unserer Studie möchten wir herausfinden, unter welchen Bedingungen Patientinnen und Patienten sowie Ärztinnen und Ärzte die Nutzung von künstlicher Intelligenz in der Gesundheitsversorgung akzeptieren.

Wie läuft unsere Studie ab?

Im Zentrum der Befragung stehen Ihre Einschätzungen zu kurzen Video-Szenarien, in denen ein möglicher Einsatz künstlicher Intelligenz in der Gesundheitsversorgung gezeigt wird. Um vergleichend Gemeinsamkeiten und Unterschiede in der Akzeptanz künstlicher Intelligenz zwischen Versicherten und ÄrztInnen identifizieren zu können, werden darüber hinaus unter anderem Fragen zur persönlichen Gesundheit, der Nutzung von Technik sowie der eigenen Persönlichkeit gestellt. Für die Betrachtung der Videos bitten wir Sie, die Studie in einer ruhigen Umgebung durchzuführen und den Ton auf Ihrem Endgerät einzuschalten.

Was ist der Nutzen Ihrer Studienteilnahme?

Mit Ihrer Teilnahme an unserer Studie leisten Sie einen wichtigen Forschungsbeitrag für die zukünftige medizinische Versorgung. Durch Ihre Einschätzungen und Anmerkungen erarbeiten wir Empfehlungen zur Gestaltung KI-basierter Lösungen in der Gesundheitsversorgung.

Sind meine Daten sicher?

Die Befragung dient ausschließlich wissenschaftlichen Zwecken und wird in Einklang mit den Regeln der DSGVO ausgewertet. Ihre Daten werden anonym erhoben und Sie haben jederzeit die Möglichkeit, die Studie abzubrechen oder Ihre Einwilligung zur Teilnahme zurückzuziehen. Setzen Sie sich hierfür bitte mit unserem Forschungsteam in Verbindung.

Wer führt die Studie durch?

Die Befragung ist Teil des vom **Gemeinsamen Bundesausschusses (G-BA)** geförderten Forschungsprojekts **Künstliche Intelligenz in der Versorgung - Bedingungen der Akzeptanz von Versicherten und ÄrztInnen (KI-BA)**. Das Projekt wird in Kooperation zwischen dem **Institut für Psychogerontologie** der **Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)** und dem **Fraunhofer-Institut für Integrierte Schaltungen (IIS)** durchgeführt.

Bei Fragen oder Anmerkungen wenden Sie sich gerne an unser Forschungsteam

Fraunhofer IIS: Martina Simon

IPG (FAU): Dr. Stefan T. Kamin

Wir danken Ihnen sehr herzlich für Ihr Mitwirken an unserer Studie!

Durch einen Klick auf die Schaltfläche rechts unten können Sie die Befragung starten.

Datenschutz

Information zum Datenschutz

Wir nehmen den Schutz Ihrer persönlichen Daten sehr ernst und folgen dabei den aktuellen Vorschriften der Datenschutzgrundverordnung (DSGVO) und dem Bayerischen Datenschutzgesetz (BayDSG). Alle Angaben, die Sie im Zuge der Befragung machen, sind anonym und werden von uns streng vertraulich behandelt. Sie dienen ausschließlich zu wissenschaftlichen Zwecken und werden nicht an unbefugte Dritte weitergegeben. Die Teilnahme an der Studie ist freiwillig. Sie können Ihre Teilnahme an der Studie jederzeit abbrechen. Ihre Angaben werden automatisch und anonym mit Datum und Uhrzeit gespeichert. Sollten Sie nach Abschluss der Befragung Ihre Einwilligung widerrufen wollen, senden Sie uns bitte eine E-Mail an ipg-gerotest@fau.de, in der Sie das Datum und die Uhrzeit (möglichst genau) Ihrer Teilnahme an der Studie angeben, damit wir Ihre Daten entsprechend löschen können. Zu Ihren Rechten als Betroffene finden Sie Information unter <https://www.fau.de/fau/leitung-und-gremien/gremien-und-beauftragte/beauftragte/datenschutzbeauftragter/#betroffenenrechte>.

Einwilligung

(consent) Ich habe die Information zum Datenschutz gelesen und willige ein, an der Studie teilzunehmen. (1 Zustimmung)

Idcode

Bevor die Studie beginnt, möchten wir Sie bitten, ein persönliches Codewort zu erstellen. Dieses setzt sich aus Ihren Antworten auf die untenstehenden Fragen zusammen. Der Code wird dafür benutzt, die aktuelle Befragung zu anonymisieren, so dass Sie die Möglichkeit haben, Ihre Einwilligung zur Studie jederzeit zurückzuziehen

Es werden Umlaute (ä, ö und ü) und sowohl Groß- als auch Kleinschreibung akzeptiert. Beim letzten Item geben Sie bitte zwei Ziffern ein.

(idcode_1) dritte Buchstabe Ihres Vornamens (z.B. Peter) (idcode_2)

erster Buchstabe des Vornames Ihrer Mutter (z.B. Maria)

(idcode_3) dritter Buchstabe des Vornamens Ihrer Mutter (z.B. Maria) (idcode_4)

erster Buchstabe des Geburtsnamens Ihrer Mutter (z.B. Meier) (idcode_5) dritter

Buchstabe des Geburtsnamens Ihrer Mutter (z.B. Meier) (idcode_6) erster Buchstabe Ihrer Geburtsstadt (z.B. Berlin)

(idcode_7) dritter Buchstabe Ihrer Geburtsstadt (z.B. Berlin)

(idcode_8) der Tag, an dem Sie geboren wurden? (z. B. 07, wenn Sie am 7. August Geburtstag haben)

Soziodemographische Faktoren

Fragen zur eigenen Person und beruflichen Situation

Vielen Dank, dass Sie an der Studie teilnehmen und sich für die Beantwortung unseres Fragebogens Zeit nehmen. Im Folgenden möchten wir Sie nun bitten, einige Angaben zu Ihrer Person und zu Ihrer beruflichen Situation zu beantworten. Diese Angaben sind für uns wichtig, um unterschiedliche Meinungen zu künstlicher Intelligenz besser verstehen zu können.

Alter und Geschlecht

Kurzbeschreibung:

Geburtsjahr, biologisches Geschlecht sowie Geschlechtsidentifikation.

Items:

(byear) In welchem Jahr wurden Sie geboren? Bitte geben Sie das vierstellige Geburtsjahr an (z.B. 1970).
_____ (numerisch, vierstellig)

(sex) Welches Geschlecht wurde bei Ihrer Geburt in Ihre Geburtsurkunde eingetragen?

- 0 weiblich
- 1 männlich

(gender) Welchem Geschlecht ordnen Sie sich selbst zu?

- 0 Weiblich
- 1 Männlich
- 2 Transgeschlechtlich
- 3 Anders Geschlecht und zwar: _____ (gender_txt)
- 4 Keine Angabe

Kommentar:

Soziodemographische Daten werden analog zu verschiedenen bevölkerungsrepräsentativen Stichproben (SHARE, DEAS, SOEP) erfasst. Die Geschlechtskategorie „divers“ wird nicht erfasst, da dieser Eintrag in das Geburtenregister erst ab 2017 möglich ist.

Quellen:

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., . . . Zuber, S. (2013). Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *International Journal of Epidemiology*, 42(4), 992–1001. doi:10.1093/ije/dyt088

Kantar Public. 2020. SOEP-Core – 2019: Personenfragebogen, Stichprobe Q (LGB). SOEP Survey Papers 929:Series A. Berlin: DIW/SOEP

Räumlicher Kontext

Gemeindegröße

Kurzbeschreibung:

Erfassung der Stadt- und Gemeindetypen.

Items:

(area_doc) Wo befindet sich Ihre Praxis / Ihr Arbeitsort?

- 1 in einer Metropole (über 500.000 EinwohnerInnen)
- 2 in einer Großstadt (über 100.000 EinwohnerInnen)
- 3 am Rand oder in einem Vorort einer Großstadt
- 4 in einer mittelgroßen Stadt (40.000 – 100.000 EinwohnerInnen)
- 5 in einer Kleinstadt (5.000 – unter 40.000 EinwohnerInnen)
- 6 in einem Dorf
- 7 in einem alleinstehenden Haus oder Hof am Land
- 8 keine Angabe

Kommentar:

Stadt- und Gemeindeklassifikation in Anlehnung an Gemeindereferenz des Bundesinstituts für Bau-, Stadt- und Raumforschung (BBSR); abgewandelt zum Dienstort (Wohnort im Original)

Quellen:

www.bbsr.bund.de

Postleitzahl

Kurzbeschreibung:

Erfassung der Postleitzahl.

Items:

(plz) Bitte geben Sie die Postleitzahl Ihrer Praxis / Ihres Arbeitsorts an. Diese Information hilft uns, regionale Unterschiede zu berücksichtigen.

Meine Postleitzahl ist: _____ (numerisch, fünfstellig)

Kommentar:

Wichtig für Differenzierung ländlicher/urbaner Raum sowie Nachrekrutierung; eigene Items

Beruf

Facharzt & Ausbildung

Kurzbeschreibung:

Facharzt und Arbeitssituation

Items:

Bitte kreuzen Sie Zutreffendes an:

- (spec_1) Kein/e Fachärztin / Facharzt (spec_2)
Innere Medizin
- (spec_3) Allgemeinmedizin
- (spec_4) Chirurgie (spec_5)
Anästhesiologie
- (spec_6) Frauenheilkunde / Geburtshilfe (spec_7)
Kinder- und Jugendmedizin (spec_8) Psychiatrie
und Psychotherapie (spec_9) Radiologie
- (spec_10) Augenheilkunde
- (spec_11) Neurologie
- (spec_12) Hals-Nasen-Ohrenheilkunde
- (spec_13) Urologie
- (spec_14) Haut- und Geschlechtskrankheiten
- (spec_15) Anderer Bereich

Skalierung:

	Ich bin Facharzt in folgender Fachrichtung:	Ich habe den Großteil meiner beruflichen Laufbahn in dieser Fachrichtung gearbeitet:	Aktuell arbeite ich in dieser Fachrichtung:
0	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
3	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
4	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
5	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
6	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
7	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Kommentar:

Soziodemographische Daten werden analog zu verschiedenen bevölkerungsrepräsentativen Stichproben (SHARE, DEAS, SOEP) erfasst. Mögliche Kodierung kann nach ISCED 2011 erfolgen.

Quellen:

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., . . . Zuber, S. (2013). Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *International Journal of Epidemiology*, 42(4), 992–1001. doi:10.1093/ije/dyt088

(phd) Haben Sie promoviert?

- 0 Nein
- 1 Ich promoviere zur Zeit.
- 2 Ja, und zwar im Jahr: _____ (phd_txt)

Kommentar:

Eigenes Item

Berufliche Position

Kurzbeschreibung:

Berufliche Position

Items:

(jobtitle) Bitte kreuzen Sie Zutreffendes an: Ich bin ...

- 1 Ärztin / Arzt in der Weiterbildung
- 2 Fachärztin / -arzt
- 3 Oberärztin / -arzt
- 4 Stellv. Chefärztin / -arzt
- 5 Chefärztin / -arzt
- 6 Selbständige/r Ärztin / Arzt bzw. Teilhaber/in in einer Gemeinschaftspraxis
- 7 Angestellte/r Ärztin / Arzt in einer (Gemeinschafts-)Praxis
- 8 In anderer Position beschäftigt: _____ (jobtitle_txt)

Kommentar:

Angelehnt an die untenstehende Quelle, einzelne Antwortoptionen wurden ergänzt um Selbstständige mit zu erfassen

Quellen:

Mitgliederbefragung MB-Monitor 2022 des Marburger Bundes. Online-Befragung durch Institut für Qualitätsmessung und Evaluation (IQME) durchgeführt. https://www.marburgerbund.de/sites/default/files/files/2022-08/4%20-%20Gesamtauswertung%20MB%20Monitor%202022%20-%20Grafische%20Darstellung_0.pdf

Beschäftigung

Kurzbeschreibung:

Beschäftigungsweise der ÄrztInnen unterschieden in Teil- und Vollzeit

Items:

(**jobtime**) In welchem Beschäftigungsverhältnis befinden Sie sich derzeit?

- 1 Teilzeit
- 2 Vollzeit

Quellen:

Mitgliederbefragung MB-Monitor 2022 des Marburger Bundes. Online-Befragung durch Institut für Qualitätsmessung und Evaluation (IQME) durchgeführt. https://www.marburgerbund.de/sites/default/files/files/2022-08/4%20-%20Gesamtauswertung%20MB%20Monitor%202022%20-%20Grafische%20Darstellung_0.pdf

Kurzbeschreibung:

Erfassung Privat- oder Kassenpatienten

Items:

(**ins_A**) Welche Patienten behandeln Sie?

- 1 Ich behandle ausschließlich Privatpatienten / Selbstzahler.
- 2 Ich behandle ausschließlich Kassenpatienten.
- 3 Ich behandle beide Patientengruppen.

Kommentar:

eigenes Item

Arbeitsbedingungen

Kurzbeschreibung:

Arbeitsbedingungen der ÄrztInnen

Items:

(jobcond) Wie beurteilen Sie Ihre derzeitigen Arbeitsbedingungen?

- 1 sehr schlecht
- 2 schlecht
- 3 mittelmäßig
- 4 gut
- 5 sehr gut

Quellen:

Mitgliederbefragung MB-Monitor 2022 des Marburger Bundes. Online-Befragung durch Institut für Qualitätsmessung und Evaluation (IQME) durchgeführt. https://www.marburgerbund.de/sites/default/files/files/2022-08/4%20-%20Gesamtauswertung%20MB%20Monitor%202022%20-%20Grafische%20Darstellung_0.pdf

Arbeitsbelastung

Kurzbeschreibung:

Arbeitsbelastung der ÄrztInnen

Items:

(jobstress) Wie schätzen Sie Ihre Arbeitsbelastung ein?

- 1 Ich empfinde bei meiner Arbeit keinen Stress.
- 2 Der Stress hält sich in Grenzen.
- 3 Ich bin häufig überlastet.
- 4 Ich gehe ständig über meine Grenze.

Quellen:

Mitgliederbefragung MB-Monitor 2019 des Marburger Bundes. Online-Befragung durch Institut für Qualitätsmessung und Evaluation (IQME) durchgeführt. https://www.marburgerbund.de/sites/default/files/files/2020-01/MB-Monitor-2019_grafische-Gesamtauswertung.pdf

Zeit für PatientInnen

Kurzbeschreibung:

Zeit für PatientInnen

Items:

(pattime) Steht Ihnen für die Behandlung Ihrer Patienten ausreichend Zeit zur Verfügung?

- 0 nein
- 1 ja

Quellen:

Mitgliederbefragung MB-Monitor 2017 des Marburger Bundes. Online-Befragung durch Institut für Qualitätsmessung und Evaluation (IQME) durchgeführt. <https://www.marburger-bund.de/sites/default/files/files/2018-09/mb-monitor-2017-grafische-darstellung.pdf>

Arzt-Patienten-Beziehung

Kurzbeschreibung:

Arzt-Patienten-Beziehung aus ärztlicher Sichtweise

Items:

Bitte schätzen Sie für jede der folgenden Aussagen ein, wie gut diese auf Sie zutrifft.

- (patrel_1) Ich nehme meine Patientinnen und Patienten so an, wie sie sind.
- (patrel_2) Ich berücksichtige die individuellen Lebensumstände und Bedürfnisse meiner Patientinnen und Patienten.
- (patrel_3) Bei meinen Entscheidungen berücksichtige ich den Kostenfaktor für mich / meinen Arbeitgeber.
- (patrel_4) Bei meinen Entscheidungen berücksichtige ich den Kostenfaktor für meine Patientin oder meinen Patient.
- (patrel_5) Ich treffe Entscheidungen immer gemeinsam mit meiner Patientin oder meinem Patient. (patrel_6) Ich treffe medizinische Entscheidungen basierend auf dem aktuellen Stand der Evidenz.

Skalierung:

- 1 trifft überhaupt nicht zu
- 2 trifft eher nicht zu
- 3 weder noch
- 4 trifft eher zu
- 5 trifft voll und ganz zu

Kommentar:

Eigenes Item; angelehnt an <https://www.aekb.de/fileadmin//mitgliederzeitschrift/2015/b1503.pdf>

Berufswechsel

Kurzbeschreibung:

Berufswechsel

Items:

(jobchange) Erwägen Sie, Ihre ärztliche Tätigkeit ganz aufzugeben?

- 0 nein
- 1 ja, und zwar weil: _____ (jobchange_txt)
- 2 weiß nicht

Quellen:

Mitgliederbefragung MB-Monitor 2022 des Marburger Bundes. Online-Befragung durch Institut für Qualitätsmessung und Evaluation (IQME) durchgeführt. https://www.marburgerbund.de/sites/default/files/files/2022-08/4%20-%20Gesamtauswertung%20MB%20Monitor%202022%20-%20Grafische%20Darstellung_0.pdf

Digitalisierung

Kurzbeschreibung:

Beurteilung des Digitalisierungsstands der Einrichtung

Items:

(digi) Wie beurteilen Sie den Grad der Digitalisierung in Ihrer Einrichtung / Praxis?

- 1 gering
- 2 eher gering
- 3 eher hoch
- 4 sehr hoch

Quellen:

Mitgliederbefragung MB-Monitor 2022 des Marburger Bundes. Online-Befragung durch Institut für Qualitätsmessung und Evaluation (IQME) durchgeführt. https://www.marburgerbund.de/sites/default/files/files/2022-08/4%20-%20Gesamtauswertung%20MB%20Monitor%202022%20-%20Grafische%20Darstellung_0.pdf

Schulungen

Kurzbeschreibung:

IT-Schulungen

Items:

(train) Finden regelmäßig Schulungen für IT-gestützte Arbeitsabläufe statt?

- 0 nein
- 1 ja
- 2 weiß nicht

Quellen:

Mitgliederbefragung MB-Monitor 2022 des Marburger Bundes. Online-Befragung durch Institut für Qualitätsmessung und Evaluation (IQME) durchgeführt. https://www.marburgerbund.de/sites/default/files/files/2022-08/4%20-%20Gesamtauswertung%20MB%20Monitor%202022%20-%20Grafische%20Darstellung_0.pdf

Ausstattung

Kurzbeschreibung:

Verwendung / Vorhandensein von Technologien am Arbeitsort

Items:

Bitte geben Sie an, inwieweit die folgenden Technologien an Ihrem Arbeitsort verwendet werden:

- (jobdigi_1) WLAN für PatientInnen
- (jobdigi_2) Tablet-gestützte PatientInnen-Aufnahme / Anamneses
- (jobdigi_3) Digitale Aufklärungsbögen
- (jobdigi_4) Digitale Rezepte (jobdigi_5)
Digitale Überweisungen
- (jobdigi_6) Digitale Krankschreibung / AU
- (jobdigi_7) Digitale Krankenakte
- (jobdigi_8) Digitale Gesundheitsanwendungen (DiGA, gemäß §139e SGB V) (jobdigi_9)
Befundübermittlung per E-Mail
- (jobdigi_10) Rechnungsübermittlung per E-Mail
- (jobdigi_11) Telemedizin / Video-Konsultation von Fachleuten (jobdigi_12)
Online-Terminvereinbarung
- (jobdigi_13) Online-Sprechstunde
- (jobdigi_14) Augmented Reality / Virtual Reality, z.B. für Trainingszwecke oder OPs
- (jobdigi_15) Künstliche Intelligenz, z.B. bei der Auswertung bildgebender Verfahren
- (jobdigi_16) Roboter-gestützte Behandlungen, z.B. OPs

- 1 nutzen wir
- 2 nutzen wir nicht
- 3 in meinem Arbeitsbereich nicht relevant

Quellen:

Items in Anlehnung an die BitkomUmfrage, angepasst & erweitert:
<https://www.bitkom.org/Presse/Presseinformation/Digitalisierung-Medizin-2022#item-16631-close>

KI-Erfahrungen

Kurzbeschreibung:

Berufliche Erfahrung mit KI

Items:

(jobexp) Haben Sie Erfahrungen mit dem Einsatz künstlicher Intelligenz in Ihrem Beruf?

Skalierung:

0 Nein

1 Ja, und zwar: _____ (jobexp_txt)

Kommentar:

Eigenes Item

Technikmotivation

Kurzbeschreibung:

Erfassung der persönlichen Technikmotivation.

Items:

Wie ist Ihre Haltung zu moderner Technik?

Nr.		Dimension
stai_1	Ich strenge mich so lange an, bis ein neues Gerät funktioniert, wie ich es will.	TGE
stai_2	Ich vertraue moderner Technik.	PST
stai_3	Die Nutzung moderner Technik hilft mir, wichtige Entscheidungen zu treffen.	PAU
stai_4	Ich übe so lange mit einem neuen Gerät, bis ich dieses optimal benutzen kann.	TGE
stai_5	Technischen Neuerungen sehe ich mit Zuversicht entgegen.	PST
stai_6	Die Nutzung moderner Technik hilft mir bei der Bewältigung des Alltags.	PAU
stai_7	Ich verstärke meine Anstrengung, wenn ein neues Gerät schwieriger zu bedienen ist als erwartet.	
stai_8	Ich vertraue darauf, dass neue technische Innovationen hohen Sicherheitsstandards genügen.	PST
stai_9	Die Nutzung moderner Technik hilft mir, mein Leben unabhängig zu führen.	PAU
stai_10	Wenn ein neues technisches Gerät nicht so funktioniert wie ich es will, spornt mich das zu mehr Anstrengung an.	
stai_11	Moderne Technik gibt mir ein Gefühl der Sicherheit.	PST
stai_12	Die Nutzung moderner Technik hilft mir, meine täglichen Aufgaben effektiver zu bewältigen.	
stai_13	Alles in allem nutze ich moderne Technik sehr häufig.	TC
stai_14	Ich interessiere mich für technische Neuerungen.	TC
stai_15	Alles in allem schätze ich mich als kompetent im Umgang mit moderner Technik ein.	TC

Skalierung:

- 1 trifft überhaupt nicht zu
- 2 trifft eher nicht zu
- 3 weder noch
- 4 trifft eher zu
- 5 trifft voll und ganz zu

Kommentar:

Erfassung der Technikmotivation über STAI (Kamin & Lang, 2013); die Items 13-15 dienen zur Validierung und können als Outcome genutzt werden.

Quellen:

Kamin, S. T., & Lang, F. R. (2013). The Subjective Technology Adaptivity Inventory (STAI): A motivational measure of technology usage in old age. *Gerontechnology*, 12(1), 16–25.

Video – Einführung

Was ist eigentlich künstliche Intelligenz?

Im Folgenden möchten wir Sie nun bitten, sich das unten eingefügte Video anzusehen. Das Video soll Ihnen helfen, einen kurzen generellen Einblick in das Thema der künstlichen Intelligenz zu erhalten.

Das Video ist vertont, bitte schalten Sie also die Tonwiedergabe Ihres Computers oder Ihres mobilen Endgeräts ein. Sie können das Video jederzeit pausieren oder auch wiederholt ansehen. Wenn Sie das Video angesehen und verstanden haben, dann fahren Sie bitte mit der Befragung fort. Um das Video zu starten, klicken sie bitte auf den Pfeil in der Mitte des Bildes / links unten im Bildfenster.

[KI-BA_KI-Einfuehrung_mit_Sound.mp4]

KI-Akzeptanz in der Gesundheitsversorgung

Wie ist Ihre Meinung zu künstlicher Intelligenz?

Vielen Dank, dass Sie sich das Video angeschaut haben. Nun interessieren wir uns dafür, welche Meinung Sie zum Einsatz künstlicher Intelligenz in der Gesundheitsversorgung haben.

Kurzbeschreibung:

Erfassung der KI-Akzeptanz in der Gesundheitsversorgung.

Items:

Bitte beurteilen Sie die folgenden Aussagen zum Einsatz künstlicher Intelligenz in der Gesundheitsversorgung.

Nr.	Künstliche Intelligenz in der Gesundheitsversorgung ..	Pol.	Dimension
accept1_1	erhöht mein Vertrauen in medizinische Empfehlungen.	+	Vertrauen
accept1_2	verringert mein Vertrauen in medizinische Empfehlungen.	-	Vertrauen
accept1_3	finde ich zuverlässig.	+	Vertrauen
accept1_4	verringert mein Vertrauen in medizinische Diagnosen.	-	Vertrauen
accept1_5	finde ich unsicher.	-	Vertrauen
accept1_6	erfüllt mich mit Angst.	+	Emotion
accept1_7	bereitet mir Unbehagen.	+	Emotion
accept1_8	fühlt sich für mich bedrohlich an.	+	Emotion
accept1_9	ist beklemmend für mich.	+	Emotion
accept1_10	lässt mich sorgenvoll in die Zukunft blicken.	+	Emotion
	Durch den Einsatz von künstlicher Intelligenz in der Gesundheitsversorgung ..		
accept2_11	unterliegen medizinische Entscheidungen meiner Kontrolle.	+	Kontrolle
accept2_12	kann ich medizinische Entscheidungen hinterfragen.	+	Kontrolle
accept2_13	trifft der Mensch die wichtigen medizinischen Entscheidungen.	+	Kontrolle

Anlage 1: AP 3 – Codebuch Akzeptanzstudie ÄrztInnen



accept2_14 unterliegt der Umgang mit medizinischen Empfehlungen meiner Kontrolle.

+ Kontrolle

Anlage 1: AP 3 – Codebuch Akzeptanzstudie ÄrztInnen



accept2_15	habe ich Einfluss auf medizinische Empfehlungen.	+	Kontrolle
accept2_16	profitieren Patienten von maßgeschneiderten Therapien.	+	Nutzen
accept2_17	wird die Gesundheitsversorgung von Patienten verbessert.	+	Nutzen
accept2_18	können Krankheiten besser behandelt werden.	+	Nutzen
accept2_19	profitiert das Gesundheitssystem insgesamt.	+	Nutzen
accept2_20	werden medizinische Empfehlungen für Menschen verbessert.	+	Nutzen

Skalierung:

- 1 trifft überhaupt nicht zu
- 2 trifft eher nicht zu
- 3 weder noch
- 4 trifft eher zu
- 5 trifft genau zu

Kommentar:

Eigene Entwicklung einer KI-Akzeptanzskala in der Versorgung

Quellen:

Hösch et al. (in prep).

Allgemeine KI-Akzeptanz

Kurzbeschreibung:

Erfassung der allgemeinen KI-Akzeptanz.

Items:

Bitte bewerten Sie folgende Aussagen zur künstlichen Intelligenz aus Ihrer Perspektive.

(atai_1) Ich habe Angst vor künstlicher Intelligenz. (atai_2) Ich vertraue künstlicher Intelligenz.

(atai_3) Künstliche Intelligenz wird die Menschheit zerstören.

(atai_4) Künstliche Intelligenz wird eine Bereicherung für die Menschheit sein. (atai_5) Künstliche Intelligenz wird für viel Arbeitslosigkeit sorgen.

Skalierung:

0 starke Ablehnung

10 starke Zustimmung

Kommentar:

Erfassung der allgemeinen KI-Akzeptanz durch die ATAI-Skala von Skindermann et al. (2021)

Quellen:

Sindermann, C., Sha, P., Zhou, M., Wernicke, J., Schmitt, H. S., Li, M., Sariyska, R., Stavrou, M., Becker, B. & Montag, C., (2021). Assessing the Attitude Towards Artificial Intelligence: Introduction of a Short Measure in German, Chinese, and English Language. KI - Künstliche Intelligenz: Vol. 35, No. 1. Springer. (S. 109-118). DOI: 10.1007/s13218-020-00689-0

Video – Stimuli

Künstliche Intelligenz - ein mögliches Anwendungsbeispiel

Wir möchten Sie nun erneut darum bitten, sich ein kurzes Video anzusehen. Dieses mal geht es um ein mögliches Anwendungsbeispiel einer konkreten KI-basierten Lösung in der Gesundheitsversorgung.

Das Video ist vertont, bitte schalten Sie also die Tonwiedergabe Ihres Computers oder Ihres mobilen Endgeräts ein. Sie können das Video jederzeit pausieren oder auch wiederholt ansehen. Wenn Sie das Video angesehen und verstanden haben, dann fahren Sie bitte mit der Befragung fort. Um das Video zu starten, klicken Sie bitte auf den Pfeil in der Mitte des Bildes / links unten im Bildfenster.

[Randomisierte Stimuli:

01 = EKG-KI-BA_KI-K_mit_Sound.mp4
02 = EKG-KI-BA_KI-L_mit_Sound.mp4
03 = EKG-KI-BA_ME-K_mit_Sound.mp4
04 = EKG-KI-BA_ME-L_mit_Sound.mp4
05 = Insulin-KI-BA_KI-K_mit_Sound.mp4
06 = Insulin-KI-BA_KI-L_mit_Sound.mp4
07 = Insulin-KI-BA_ME-K_mit_Sound.mp4
08 = Insulin-KI-BA_ME-L_mit_Sound.mp4
09 = Darm-KI-BA_KI-K_mit_Sound.mp4
10 = Darm-KI-BA_KI-K_mit_Sound.mp4
11 = Darm-KI-BA_ME-K_mit_Sound.mp4
12 = Darm-KI-BA_ME-L_mit_Sound.mp4]

(vidop) Haben Sie Anmerkungen zu den Videos? (offenes Textfeld)

Manipulationcheck

Was ist Ihre Meinung zum gezeigten Anwendungsbeispiel?

Vielen Dank, dass Sie sich das Video angesehen haben. Nun interessieren wir uns dafür, wie Sie das gezeigte KI-basierte Anwendungsbeispiel bewerten und welche Meinung Sie dazu haben. Weiterhin ist es für uns wichtig zu erfahren, ob technische Probleme beim Abspielen des Videos aufgetreten sind und ob die Inhalte verständlich waren.

Kurzbeschreibung: Technische Evaluation

Allgemeine technische Abfrage.

Items:

Sind technische Probleme beim Abspielen des Videos aufgetreten?

(tecinfo_1) Konnten Sie das Video abspielen?

(tecinfo_2) Hat der Ton beim Abspielen des Videos funktioniert?

Skalierung:

0 nein

1 ja

Kommentar:

Eigene Items

Kurzbeschreibung: Manipulationscheck

Manipulationscheck zur Absicherung der experimentellen Bedingungen.

Items:

Bitte denken Sie nun an die KI-Anwendung, die wir Ihnen im vorherigen Video präsentiert haben. Welche der folgenden Aussagen treffen zu?

- (manip1_1) Die künstliche Intelligenz führt automatisiert die weitere Behandlung durch.
- (manip1_2) Der Arzt oder Patient kann eine eigene vom Systemvorschlag abweichende Behandlung durchführen.
- (manip1_3) Die künstliche Intelligenz vergleicht die Daten des Patienten mit den Daten anderer. Patienten.
- (manip1_4) Die künstliche Intelligenz greift im Hintergrund auf eine große klinische Datenbank zurück.

Skalierung:

- 0 nein
- 1 ja
- 2 weiß ich nicht

Kommentar:

Eigene Items;

	AI completeness low	AI completeness high
Explainability low	ME-K: 1 nein 2 ja 3 nein 4 nein	KI-K: 1 nein 2 nein 3 nein 4 nein
Explainability high	ME-L: 1 nein 2 ja 3 ja 4 ja	KI-L: 1 nein 2 nein 3 ja 4 ja

(manip2) Was war der medizinische Anwendungskontext der gezeigten KI-Anwendung?

- 1 Parkinson
- 2 Diabetes
- 3 Darmspiegelung (Koloskopie)
- 4 Zahnwurzelbehandlung
- 5 Elektrokardiogramm (EKG)
- 6 Keiner der aufgezählten
- 7 Das weiss ich nicht

Kommentar:

Eigene Items

Technikakzeptanz

Kurzbeschreibung:

Bewertung der Stimuli mit Blick auf klassische Indikatoren der Technikakzeptanzforschung.

Items:

Bitte geben Sie an, wie Sie die im Video gezeigte KI-Anwendung aus Ihrer Sicht als Ärztin oder Arzt bewerten.

(tam_1) Ich kann mir vorstellen, diese KI-Anwendung zu verwenden. (tam_2)

Wenn es nötig wäre, würde ich diese KI-Anwendung verwenden. (tam_3) Ich finde diese KI-Anwendung nützlich.

(tam_4) Diese KI-Anwendung würde vieles komfortabler machen. (tam_5)

Diese KI-Anwendung würde mir mehr Sicherheit geben. (tam_6) Diese KI-Anwendung würde meinen Arbeitsalltag erleichtern. (tam_7) Diese KI-Anwendung ist leicht verständlich.

(tam_8) Ich glaube, diese KI-Anwendung wäre leicht zu bedienen.

(tam_9) Ich denke, dass die Interaktion mit dieser KI-Anwendung klar und nachvollziehbar ist.

Skalierung:

- 1 Trifft gar nicht zu
- 2 Trifft eher nicht zu
- 3 Trifft eher zu
- 4 Trifft genau zu

Kommentar:

In Anlehnung an das Technikakzeptanzmodell; abhängige Variablen. Ursprüngliche Skalierung von Davis et al. (1989): 7-stufige Skala (likely-unlikely and anchor points extremely, quite, slightly and neither); Items 2, 5, 6, 8 & in Anlehnung an Cläßen bzw. Venkatesh

Quellen:

Cläßen, K. (2013). *Zur Psychologie von Technikakzeptanz im höheren Lebensalter: Die Rolle von Technikgenerationen* (Doctoral dissertation).

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35, 982–1003.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision sciences*, 39(2), 273-315.

Persönlichkeit

Herzlichen Dank für die Bewertung der Videos!

Nun haben wir noch einige Fragen zu verschiedenen Aspekten Ihrer Person, die uns dabei helfen, die Haltung gegenüber künstlicher Intelligenz besser zu verstehen. Dabei geht es um verschiedene Bereiche wie Persönlichkeit, Gesundheit oder auch Techniknutzung.

Wir beginnen mit Fragen zur individuellen Persönlichkeit.

Persönlichkeitsdispositionen

Kurzbeschreibung:

Kurzerfassung der Persönlichkeitsdispositionen aus dem Big Five Inventory.

Items:

Inwieweit treffen die folgenden Aussagen auf Sie zu?

Nr.	Ich bin jemand, der"	Polung	Dimension
bfi_1	gründlich arbeitet.	+	Gewissenhaftigkeit
bfi_2	kommunikativ, gesprächig ist.	+	Extraversion
bfi_3	manchmal etwas grob zu anderen ist.	-	Verträglichkeit
bfi_4	originell ist, neue Ideen einbringt.	+	Offenheit
bfi_5	sich oft Sorgen macht.	+	Neurotizismus
bfi_6	zurückhaltend ist.	-	Extraversion
bfi_7	verzeihen kann.	+	Verträglichkeit
bfi_8	eher faul ist.	-	Gewissenhaftigkeit
bfi_9	aus sich herausgehen kann, gesellig ist.	+	Extraversion
bfi_10	künstlerische Erfahrungen schätzt.	+	Offenheit
bfi_11	leicht nervös wird.	+	Neurotizismus
bfi_12	Aufgaben wirksam und effizient erledigt.	+	Gewissenhaftigkeit
bfi_13	rücksichtsvoll und freundlich mit anderen umgeht.	+	Verträglichkeit
bfi_14	eine lebhaftere Phantasie, Vorstellungen hat.	+	Offenheit
bfi_15	entspannt ist, mit Stress gut umgehen kann.	-	Neurotizismus

Skalierung:

1 trifft überhaupt nicht zu

7 trifft voll und ganz zu

Kommentar:

Das Instrument wurde basierend auf dem Big Five Inventar im Rahmen des SOEP entwickelt und eingesetzt.

Quellen:

Lang, F. R., Hahne, D., Gymbel, S., Schröpper, S., & Lutsch, K. (2005). Erfassung des kognitiven Leistungspotenzials und der "Big Five" mit Computer-Assisted-Personal-Interviewing (CAPI): zur Reliabilität und Validität zweier ultrakurzer Tests und des BFI-S. *DIW Research Notes*, 9, 2005.

Zeitperspektive

Kurzbeschreibung:

Ultrakurze Erfassung der wahrgenommenen verbleibenden Spanne an Lebenszeit.

Items:

Inwiefern treffen die folgenden Aussagen bezüglich Zeit und Zukunft auf Sie zu?

(ftps_1) Auf mich warten viele Möglichkeiten in der Zukunft. (ftps_2)

Meine Zukunft ist voller Möglichkeiten.

(ftps_3) Mit zunehmendem Alter beginne ich, die Zeit als begrenzt zu erleben.

Skalierung:

1 trifft ganz und gar nicht zu

7 trifft voll und ganz zu

Kommentar:

Die Skala mit 10 Items wurde von Rohr et al. (2017) entwickelt, die Auswahl der Items für die ultrakurze Erfassung basiert auf Gellert et al. (2012). Die 7-stufige Skalierung basiert auf der VW Studie.

Quellen:

Lang, F. R., & Carstensen, L. L. (2002). Time counts: future time perspective, goals, and social relationships. *Psychology and aging, 17*(1), 125.

Rohr, M. K., John, D. T., Fung, H. H., & Lang, F. R. (2017). A three-component model of future time perspective across adulthood. *Psychology and Aging, 32*(7), 597.

Gellert, P., Ziegelmann, J. P., Lippke, S., & Schwarzer, R. (2012). Future time perspective and health behaviors: Temporal framing of self-regulatory processes in physical exercise and dietary behaviors. *Annals of Behavioral Medicine, 43*(2), 208-218.

Obsoleszenzerleben

Kurzbeschreibung:

Erfassung der wahrgenommenen Obsoleszenz.

Items:

Alle Menschen werden älter. Was Älterwerden für den Einzelnen jedoch bedeutet, kann sehr unterschiedlich sein. Wie erleben Sie Veränderungen, die mit dem Älterwerden einhergehen können?

(obs_1) Ich habe zunehmend das Gefühl, den Anschluss an die heutige Zeit verpasst zu haben. (obs_2) Ich bin voll auf der Höhe der Zeit.

(obs_3) Ich komme mit der heutigen Lebensweise immer schlechter zurecht.

(obs_4) Für die Auffassungen der jüngeren Generation habe ich immer weniger Verständnis. (obs_5) Das Leben wird für mich immer komplizierter und schwerer zu durchschauen.

Skalierung:

- 1 trifft überhaupt nicht zu
- 2 trifft eher nicht zu
- 3 weder noch
- 4 trifft eher zu
- 5 trifft voll und ganz zu

Kommentar:

Die Items wurden in Anlehnung an Brandstädter & Wentura (1994) entwickelt und werden zusammen mit den Fragen zum subjektiven Alterserleben erhoben.

Quellen:

Brandstädter, J., & Wentura, D. (1994). Veränderungen der Zeit- und Zukunftsperspektive im Übergang zum höheren Erwachsenenalter: entwicklungspsychologische und differentielle Aspekte. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 26(1), 2–21.

Subjektives Alterserleben

Kurzbeschreibung:

Erfassung des subjektiven Alterserlebens.

Items:

Alle Menschen werden älter. Was Älterwerden für den Einzelnen jedoch bedeutet, kann sehr unterschiedlich sein. Wie erleben Sie Veränderungen, die mit dem Älterwerden einhergehen können?

(pgcms_1) Je älter ich werde, desto schlimmer wird alles. (pgcms_2) Ich habe noch genauso viel Schwung wie letztes Jahr. (pgcms_3) Je älter ich werde, desto weniger nützlich bin ich.

(pgcms_4) Mit zunehmendem Alter ist mein Leben besser, als ich erwartet habe. (pgcms_5) Ich bin jetzt genauso glücklich, wie ich es in jungen Jahren war.

Skalierung:

- 1 Trifft gar nicht zu
- 2 Trifft eher nicht zu
- 3 Trifft eher zu
- 4 Trifft genau zu

Kommentar:

Die Items stammen aus der PGCMS (Lawton, 1975) und werden in verschiedenen Panelstudien (z.B. HRS, DEAS) zur Erfassung des subjektiven Alterserlebens und des persönlichen Altersbildes genutzt. Die Items werden zusammen mit den Fragen zur Obsoleszenz erhoben.

Quellen:

Lawton, M.P. (1975). The Philadelphia Geriatric Center Morale Scale: A revision. *Journal of Gerontology*, 30, 85-89.

Novelty

Kurzbeschreibung:

Erfassung der intellektuellen Neugier und Wissberigkeit (curiosity).

Items:

Im Folgenden möchten wir gerne erfahren, wie Sie mit neuen Dingen und Erfahrungen umgehen. Bitte schätzen Sie für jede der folgenden Aussagen ein, wie gut diese auf Sie zutrifft.

(nov_1) Ich profitiere davon, Neues zu lernen.

(nov_2) Bei neuen Entwicklungen auf dem Laufenden zu bleiben, unterstützt einen unabhängigen Lebensstil im Alter.

(nov_3) Sich mit neuen Dingen zu beschäftigen, verbessert meine Ressourcen (z.B. Wissen, soziale Kontakte, ...)

(nov_4) Andere werden mich verurteilen, wenn ich mit neuen Technologien nicht zurecht komme. (nov_5) Es kostet zu viel Zeit und Energie, etwas Neues benutzen zu lernen.

(nov_6) Sich auf neue Dinge einzulassen, stört meine täglichen Routinen. (nov_7) Ich probiere gerne neue Dinge aus.

(nov_8) Ich verlasse gerne meine Komfortzone.

(nov_9) Ich bevorzuge bekannte Aktivitäten gegenüber neuen Abenteuern.

Skalierung:

6 trifft überhaupt nicht zu

7 trifft eher nicht zu

8 weder noch

9 trifft eher zu

10 trifft voll und ganz zu

Kommentar:

Eigene Items

Neugier

Kurzbeschreibung:

Erfassung der intellektuellen Neugier und Wissberigkeit (curiosity).

Items:

Wie ist Ihre persönliche Erfahrung mit Blick auf neue Situationen und Erfahrungen?

(cur_1) Ich sehe herausfordernde Situationen als eine Gelegenheit, zu lernen und mich weiterzuentwickeln.

(cur_2) Ich bin stets auf der Suche nach Erlebnissen, die die Art und Weise, wie ich über mich selbst und die Welt denke, herausfordern.

(cur_3) Ich suche Situationen auf, in denen es wahrscheinlich ist, dass ich sehr gründlich über etwas nachdenken muss.

(cur_4) Ich genieße es, etwas über Themen zu lernen, die mir nicht vertraut sind. (cur_5) Ich finde es faszinierend, neue Informationen zu lernen.

Skalierung:

1 trifft ganz und gar nicht zu 7
trifft voll und ganz zu

Kommentar:

Die Items stammen aus der curiosity Skala von Kashdan et al (2018) und wurden in der VW Studie (7- stufig) eingesetzt.

Quellen:

Kashdan, T. B., Stikma, M. C., Disabato, D. J., McKnight, P. E., Bekier, J., Kaji, J., & Lazarus, R. (2018). The Five-Dimensional Curiosity Scale: Capturing the bandwidth of curiosity and identifying four unique subgroups of curious people. *Journal of Research in Personality*, 73, 130–149.

Lebenszufriedenheit

Kurzbeschreibung:

Einschätzung der eigenen subjektiven Lebenszufriedenheit.

Items:

(lifesat) Wie zufrieden sind Sie, alles in allem, mit Ihrem Leben?

Skalierung:

0 ganz und gar unzufrieden 10

ganz und gar zufrieden

Kommentar:

Erfassung analog zur Erhebung in diversen Panelstudien

Quellen:

Engstler, H., Alcantara, A. L., Luitjens, M., & Klaus, D. (2019). German Ageing Survey, Deutscher Alterssurvey (DEAS): Documentation of instruments and variables 1996 – 2017. doi: 10.5156/DEAS.1996-2017.D.002

Gesundheitliche Ressourcen

Gesundheit

Vielen Dank für die Beantwortung der Fragen zur Persönlichkeit!

Wir glauben, dass die Haltung gegenüber künstlicher Intelligenz in der Gesundheitsversorgung auch mit gesundheitlichen Aspekten zusammenhängt. Mit den nächsten Fragen möchten wir mehr darüber erfahren.

Subjektive Gesundheit

Kurzbeschreibung:

Einschätzung der eigenen subjektiven Gesundheit

Items:

(health) Wie würden Sie Ihren gegenwärtigen Gesundheitszustand beschreiben?

- 1 schlecht
- 2 weniger gut
- 3 zufriedenstellend
- 4 gut
- 5 sehr gut

Kommentar:

Erfassung der subjektiven Gesundheit analog zur Erhebung im SOEP.

Quellen:

SOEP-IS Group, 2021. SOEP-IS 2019 – Questionnaire for the SOEP Innovation Sample. SOEP Survey Papers 956: Series A – Survey Instruments (Erhebungsinstrumente). Berlin: DIW Berlin/SOEP

Einsamkeit

Kurzbeschreibung:

Ultrakurze Erfassung von Einsamkeitserleben.

Items:

Bitte geben Sie für jede Frage an, wie oft Sie sich so fühlen. (lone_1) Wie häufig vermissen Sie Gesellschaft?

(lone_2) Wie häufig fühlen Sie sich isoliert von anderen?

(lone_3) Wie häufig fühlen Sie sich einsam?

Skalierung:

- 1 niemals
- 2 selten
- 3 manchmal
- 4 häufig

Kommentar:

Basiert auf Kurzversion der R-UCLA, eingesetzt in SHARE.

Quellen:

Hughes, M. E., Waite, L. J., Hawkey, L. C., & Cacioppo, J. T. (2004). A short scale for measuring loneliness in large surveys: Results from two population-based studies. *Research on aging, 26*(6), 655- 672.

Ängstlichkeit und Depression

Kurzbeschreibung:

Ultrakurze Erfassung depressiver Symptome und generalisierter Ängstlichkeit.

Items:

Jetzt geht es um die letzten 2 Wochen. Wie oft fühlten Sie sich da durch die folgenden Beschwerden beeinträchtigt?

(anx_1) Wenig Interesse oder Freude an Ihren Tätigkeiten

(anx_2) Niedergeschlagenheit, Schwermut oder Hoffnungslosigkeit (anx_3)
Nervosität, Ängstlichkeit oder Anspannung

(anx_4) Nicht in der Lage sein, Sorgen zu stoppen oder zu kontrollieren

Skalierung:

- 1 überhaupt nicht
- 2 an einzelnen Tagen
- 3 mehr als die Hälfte der Tage
- 4 beinahe jeden Tag

Kommentar:

Ultrakurzform des Gesundheitsfragebogens für Patienten (PHQ-4) zur Erfassung der diagnostischen Kernkriterien für depressive Störungen und für die Generalisierte Angststörung nach DSM-IV.

Quellen:

B. Löwe, I. Wahl, M. Rose, C. Spitzer, H. Glaesmer, K. Wingenfeld, A. Schneider, E. Brähler: A 4- item measure of depression and anxiety: Validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. In: *Journal of Affective Disorders*. 122, 2010, S. 86- 95.

Gesundheitsverhalten

Vorsorge

Kurzbeschreibung:

Erfassung des gesundheitsbezogenen Vorsorgeverhaltens.

Items:

(prev) Ich nehme regelmäßig an Vorsorge-Untersuchungen (z.B. Gesundheits-Check-Up, Krebsfrüherkennung, Zahnkontrolle) teil.

Skalierung:

- 1 Trifft überhaupt nicht zu
- 2 Trifft eher nicht zu
- 3 Trifft eher zu
- 4 Trifft voll und ganz zu

Kommentar:

Gesundheitsbezogenes Vorsorgeverhalten; eingesetzt im DEAS.

Quellen:

Bellach, B. M., Knopf, H., & Thefeld, W. (1998). The German Health Survey. 1997/98]. Gesundheitswesen (Bundesverband der Ärzte des Öffentlichen Gesundheitsdienstes (Germany)), 60, S59.

Soziale Beziehungen

Sozialen Beziehungen

Soziale Kontakte und Beziehungen zu anderen Menschen können vollkommen unterschiedlich aussehen und eine wichtige Rolle spielen, wenn es um die Bewertung von Angeboten in der Gesundheitsversorgung geht. Ihre Angaben zu den folgenden Fragen helfen uns dabei, dies genauer zu verstehen.

Zufriedenheit mit Beziehungen

Kurzbeschreibung:

Erfasst wird die globale Zufriedenheit mit Familienbeziehungen und Freunden.

Items:

(famsat) Wie zufrieden sind Sie mit Ihren Familienbeziehungen? (friendsat)
Wie zufrieden sind Sie mit Ihrem Freundeskreis?

Skalierung:

1 sehr unzufrieden

7 sehr zufrieden

Kommentar:

Items angelehnt an Panelstudien (SOEP, NRW80+, DEAS) sowie an VW Studie.

Quellen:

Lang, F. R., Lessenich, S., & Rothermund, K. (2022). *Altern als Zukunft–eine Studie der Volkswagenstiftung* (p. 173). Springer Nature.

Soziale Unterstützung

Kurzbeschreibung:

Erfasst wird die globale erhaltene Unterstützung von Familienbeziehungen und Freunden.

Items:

(famsupp) Wie häufig erhalten Sie Unterstützung und Hilfe von Ihrer Familie? (friendsupp) Wie häufig erhalten Sie Unterstützung und Hilfe aus Ihrem Freundeskreis?

Skalierung:

1 sehr selten

7 sehr häufig

Kommentar:

Items angelehnt an Panelstudien (SOEP, NRW80+, DEAS), in diesen allerdings erfasst im Rahmen von Netzwerkgeneratoren.

Quellen:

Engstler, H., Alcantara, A. L., Luitjens, M., & Klaus, D. (2019). German Ageing Survey, Deutscher Alterssurvey (DEAS): Documentation of instruments and variables 1996 – 2017. doi: 10.5156/DEAS.1996-2017.D.002

Anzahl wichtiger Beziehungen

Kurzbeschreibung:

Erfasst wird die Anzahl bedeutsamer Beziehungspartner.

Items:

(npeople) Wenn Sie nun insgesamt an die Menschen denken, die Ihnen wichtig sind und mit denen Sie in Kontakt stehen – um wieviele Personen handelt es sich dabei?

Anzahl der mir wichtigen Personen: _____ (numerisch)

Kommentar:

Indikator sozialer Ressourcenstrukturen und Eingebundenheit; Frage angelehnt an Panelstudien, häufig Eröffnung für Netzwerkgenerator (SOEP, NRW80+, DEAS).

Quellen:

Malter, F. and Börsch-Supan, A.(Eds.)(2013). SHARE Wave 4: Innovations & Methodology.
Munich: MEA, Max Planck Institute for Social Law and Social Policy.

Abschluss der Befragung

Herzlichen Dank!

Sie haben alle Fragen beantwortet, vielen Dank für Ihre Mühe! Zum Abschluss der Befragung möchten wir gerne sicherstellen, dass wir Ihre Angaben für die wissenschaftliche Auswertung verwenden können.

Reluktanz

Kurzbeschreibung:

Seriousness Checks zur Verbesserung der Datenqualität – Reluktanz beim Ausfüllen.

Items:

(scfun) Jetzt mal Hand aufs Herz: Haben Sie gerne an dieser Studie teilgenommen? Bitte antworten Sie ehrlich - Ihre Antwort hat keine Konsequenzen für Sie!

Skalierung:

- 1 Nein
- 2 Eher Nein
- 3 Eher Ja
- 4 Ja

Kommentar:

Die Items ermöglichen eine Einschätzung der Datenqualität; Anlehnung an Aust et al. (2013); standardmäßig implementiert in SoSci.

Quellen:

Aust, F., Diedenhofen, B., Ullrich, S. *et al.* Seriousness checks are useful to improve data validity in online research. *Behav Res* **45**, 527–535 (2013). <https://doi.org/10.3758/s13428-012-0265-2>

Zuverlässigkeit

Kurzbeschreibung:

Seriousness Checks zur Verbesserung der Datenqualität – Zuverlässigkeit der Angaben.

Items:

(mresp) Haben Sie alle Aufgaben so ausgeführt, wie in den jeweiligen Instruktionen gebeten? Bitte antworten Sie auch hier ehrlich - Auch diese Antwort hat für Sie keine Konsequenzen!

Skalierung:

- 1 Ich habe alle Aufgaben, wie in den Instruktionen verlangt, bewältigt.
- 2 __Manchmal__ habe ich irgendetwas geklickt, weil ich unmotiviert war oder mich einfach nicht ausgekannt habe.
- 3 Ich habe __häufig__ irgendetwas angeklickt, damit ich schnell fertig werde.

Kommentar:

Die Items ermöglichen eine Einschätzung der Datenqualität; Anlehnung an Aust et al. (2013); standardmäßig implementiert in SoSci.

Quellen:

Aust, F., Diedenhofen, B., Ullrich, S. *et al.* Seriousness checks are useful to improve data validity in online research. *Behav Res* **45**, 527–535 (2013). <https://doi.org/10.3758/s13428-012-0265-2>

Ablenkung

Kurzbeschreibung:

Seriousness Checks zur Verbesserung der Datenqualität – Ablenkung beim Ausfüllen.

Items:

(distr) Konnten Sie den Fragebogen in einem ausfüllen, ohne abgelenkt zu werden? Bitte antworten Sie auch hier ehrlich.

Skalierung:

- 1 Ja, ich habe der Studie vollste Aufmerksamkeit geschenkt und sie in einem durch gemacht.
- 2 Ich wurde durch meine Umwelt (Personen, Geräusche, etc.) __einmal__ kurz abgelenkt.
- 3 Ich wurde durch meine Umwelt (Personen, Geräusche, etc.) __mehrmals__ abgelenkt.

Kommentar:

Die Items ermöglichen eine Einschätzung der Datenqualität; Anlehnung an Aust et al. (2013); standardmäßig implementiert in SoSci.

Quellen:

Aust, F., Diedenhofen, B., Ullrich, S. *et al.* Seriousness checks are useful to improve data validity in online research. *Behav Res* **45**, 527–535 (2013). <https://doi.org/10.3758/s13428-012-0265-2>

Anmerkungen

(notes) Haben Sie noch Anmerkungen zur Studie? (offenes Textfeld)

Kontakt

Kurzbeschreibung:

Anonymisierte Teilnahme an Verlosung und Bereitschaft für weitere Studien kontaktiert zu werden.

Items:

(contact) Nochmals vielen Dank für Ihre Teilnahme an unserer Studie!

Weiterhin würden wir uns freuen, wenn wir Sie in Zukunft für weitere Studien kontaktieren dürften.

Bitte beachten Sie, dass Ihre Kontaktdaten getrennt von den Befragungsdaten gespeichert werden. Die Anonymität Ihrer Angaben ist gewährleistet.

- 1 Auswahloption Information
- 2 Auswahloption weitere Studien

Kommentar:

Spezieller Fragetyp „Kontaktdaten getrennt erheben“ zur Wahrung der Anonymität der Teilnehmer.

Quellen:

<https://www.socisurvey.de/help/doku.php/de:create:questions:contact>

KI-BA:

Künstliche Intelligenz in der Versorgung – Bedingungen der Akzeptanz von Versicherten und ÄrztInnen

**Manual Hauptstudie – Versicherte Stand:
27.09.2022**

Projektmitarbeitende:

A. Beyer, N. M. Degen, H. S. Hösch, F. R. Lang, & S. T. Kamin

<u>FORSCHUNGSDESIGN</u>	4
<u>EINLEITUNGSTEXT</u>	5
<u>DATENSCHUTZ</u>	7
<u>IDCODE</u>	8
<u>SOZIODEMOGRAPHISCHE FAKTOREN</u>	9
ALTER UND GESCHLECHT	9
HÖCHSTER BILDUNGSABSCHLUSS	10
ERWERBSTÄTIGKEIT	11
HAUSHALTSGRÖßE	12
<u>RÄUMLICHER KONTEXT</u>	13
GEMEINDEGRÖßE	13
ERREICHBARKEIT VERSORGUNGSANGEBOTE	14
POSTLEITZAHL	15
<u>VIDEO – EINFÜHRUNG</u>	16
KI-AKZEPTANZ IN DER GESUNDHEITSVERSORGUNG	17
ALLGEMEINE KI-AKZEPTANZ	19
<u>VIDEO – STIMULI</u>	20
MANIPULATIONSCHECK	21
TECHNIKAKZEPTANZ	22
USER EXPERIENCE	23
<u>PERSÖNLICHKEIT</u>	24
PERSÖNLICHKEITSDISPOSITIONEN	24
ZEITPERSPEKTIVE	26
OBSOLESZENZERLEBEN	27
SUBJEKTIVES ALTERSERLEBEN	28
NOVELTY	29
NEUGIER	30
LEBENSZUFRIEDENHEIT	31
<u>GESUNDHEITLICHE RESSOURCEN</u>	32
SUBJEKTIVE GESUNDHEIT	32
COVID-19 INFektion	33
FUNKTIONALE BEEINTRÄCHTIGUNGEN	34

EINSAMKEIT	35
ÄNGSTLICHKEIT UND DEPRESSION	36
DIAGNOSEN	37
<u>GESUNDHEITSVERHALTEN</u>	38
KRANKENHAUSAUFENTHALTE	38
ARZTKONTAKTE	39
VORSORGE	40
KRANKENVERSICHERUNG	41
ARZT-PATIENTEN-BEZIEHUNG	42
<u>SOZIALE BEZIEHUNGEN</u>	43
ZUFRIEDENHEIT MIT BEZIEHUNGEN	43
SOZIALE UNTERSTÜTZUNG	44
ANZAHL WICHTIGER BEZIEHUNGEN	45
<u>DIGITALISIERUNG UND TECHNIK</u>	46
INTERNETNUTZUNGSVERHALTEN	46
TECHNIKMOTIVATION	47
DIGITALKOMPETENZ	49
<u>ABSCHLUSS DER BEFRAGUNG</u>	50
RELUKTANZ	50
ZUVERLÄSSIGKEIT	51
ABLENKUNG	52
VERLOSUNG UND KONTAKT	53

Forschungsdesign

Kurzbeschreibung:

Grundanliegen des Vorhabens ist es zu verstehen, wie die Akzeptanz gegenüber KI-Systemen bei Versicherten und ÄrztInnen ausgeprägt ist und welche Einflussfaktoren die individuelle Akzeptanz beeinflussen. Dafür werden den Probandinnen und Probanden im Rahmen eines 2x2-faktoriellen Designs randomisierte KI-Anwendungsszenarien in Form animierter Kurzvideos präsentiert. Die Videos repräsentieren die zuvor entwickelte Kategorisierung und sind mit Blick auf die individuelle Akzeptanz gegenüber der entsprechenden KI-Anwendung zu bewerten.

Versuchsplan:

Der mehrfaktorielle Versuchsplan beinhaltet zwei unabhängige Variablen mit jeweils zwei Faktorstufen und eine unabhängige Variable mit 3 Faktorstufen (2x2x3 Design):

- IV1: AI-Explainability (high/low), d.h. die Funktionsweise der KI wird dargestellt – oder nicht
- IV2: AI-Completeness (high/low), d.h. die KI entscheidet autonom – oder der Mensch
- IV3: KI Use-case (Diabetes, EKG, Koloskopie)

Aufbau der Befragung:

Der Aufbau der Befragung ist in Abb. 1 dargestellt. Die erste Seite umfasst die Erhebung grundlegender soziodemographischer Informationen. Danach werden die Probanden gebeten, sich ein Video zur Einführung in die KI-Thematik anzuschauen. Nach diesem Video erfolgt die Erhebung der KI- Akzeptanzskala, die im Rahmen des Projektes entwickelt wurde, um interindividuelle Unterschiede in der Bewertung der KI-Lösungen zu erklären. Danach erfolgt die randomisierte Zuteilung zur einer between- subject Bedingung, innerhalb derer eine randomisierte Videosequenz zu evaluieren ist. Die Evaluation erfolgt anhand klassischer Instrumente aus der Technikakzeptanz- sowie der UX-Forschung, die gleichzeitig als zentrale abhängige Variablen fungieren. Nach der Evaluationen werden die Probanden gebeten, weitere Fragen zu potentiellen Einflussfaktoren der KI-Akzeptanz zu beantworten. Dabei handelt es sich um die Bereiche:

- Räumlicher Kontext (z.B. Versorgungssituation, Gemeindegröße)
- Soziale Beziehungen (z.B. Anzahl Beziehungspartner, soziale Unterstützung)
- Gesundheitliche Ressourcen (z.B. Krankheiten, Alltagsfunktionalität)
- Gesundheitsverhalten (z.B. Vorsorgeverhalten, Arzt-Patienten-Beziehung)
- Persönlichkeit (z.B. Zeitperspektive, Alterserleben)
- Technik und Digitalisierung (z.B. Technikmotivation, Digitalkompetenz)

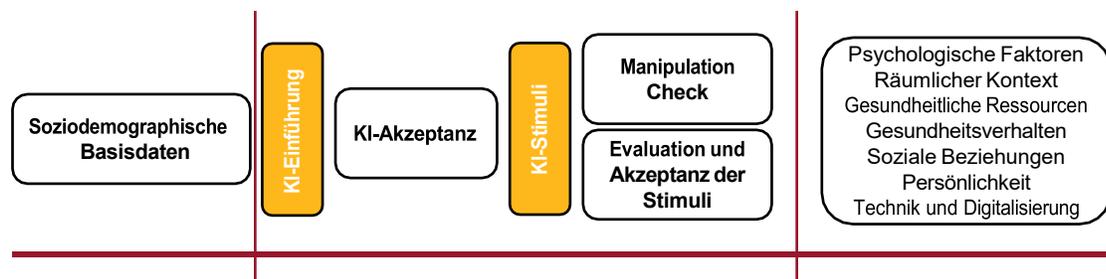


Abbildung 1: Aufbau der Befragung

Im Folgenden sind Variablennamen rot und Variablenwerte grün gekennzeichnet. Textblöcke aus dem Fragebogen sind grau hinterlegt.

Einleitungstext

Herzlich willkommen zur Online-Befragung des Projekts "KI-BA"

"Künstliche Intelligenz (KI) in der medizinischen Versorgung - Bedingungen der Akzeptanz von Versicherten und ÄrztInnen"

Was ist das Ziel unserer Studie?

Mit Hilfe künstlicher Intelligenz können in der Gesundheitsversorgung besonders komplexe Abläufe und Entscheidungsprozesse optimiert werden. Unklar ist allerdings bisher, in welcher Form diese Technologie im Gesundheitsbereich Anwendung finden kann. Viele Faktoren beeinflussen die Umsetzung und Nutzung künstlicher Intelligenz, nicht zuletzt spielt die Akzeptanz aller Beteiligten eine wesentliche Rolle.

Mit Hilfe unserer Studie möchten wir herausfinden, unter welchen Bedingungen Patientinnen und Patienten sowie Ärztinnen und Ärzte die Nutzung von künstlicher Intelligenz in der Gesundheitsversorgung akzeptieren.

Wie läuft unsere Studie ab?

Im Zentrum der Befragung stehen Ihre Einschätzungen zu kurzen Video-Szenarien, in denen ein möglicher Einsatz künstlicher Intelligenz in der Gesundheitsversorgung gezeigt wird. Für die Beantwortung des Online-Fragebogens benötigen Sie in etwa 25 Minuten. Für die Betrachtung der Videos bitten wir Sie, die Studie in einer ruhigen Umgebung durchzuführen und den Ton auf Ihrem Endgerät einzuschalten.

Was ist der Nutzen Ihrer Studienteilnahme?

Mit Ihrer Teilnahme an unserer Studie leisten Sie einen wichtigen Forschungsbeitrag für die zukünftige medizinische Versorgung. Durch Ihre Einschätzungen und Anmerkungen erarbeiten wir Empfehlungen zur Gestaltung KI-basierter Lösungen in der Gesundheitsversorgung. Unter allen Teilnehmenden verlosen wir **100 Gutscheine im Wert von je 20 Euro** (Teilnahme bis zum 30.11.2022).

Sind meine Daten sicher?

Die Befragung dient ausschließlich wissenschaftlichen Zwecken und wird in Einklang mit den Regeln der DSGVO ausgewertet. Ihre Daten werden anonym erhoben und Sie haben jederzeit die Möglichkeit, die Studie abubrechen oder Ihre Einwilligung zur Teilnahme zurückzuziehen. Setzen Sie sich hierfür bitte mit unserem Forschungsteam in Verbindung.

Wer führt die Studie durch?

Die Befragung ist Teil des vom **Gemeinsamen Bundesausschusses (G-BA)** geförderten Forschungsprojekts **Künstliche Intelligenz in der Versorgung - Bedingungen der Akzeptanz von Versicherten und ÄrztInnen (KI-BA)**. Das Projekt wird in Kooperation zwischen dem **Institut für Psychogerontologie** der **Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)** und dem **Fraunhofer Institut für Integrierte Schaltungen (IIS)** durchgeführt.

Bei Fragen oder Anmerkungen wenden Sie sich gerne an unser Forschungsteam



Helena S. Hösch, M.Sc.

Nora M. Degen, M.Sc.

Dr. Stefan T. Kamin

Wir danken Ihnen sehr herzlich für Ihr Mitwirken an unserer Studie!

Durch einen Klick auf die Schaltfläche rechts unten können Sie die Befragung starten.

Datenschutz

Information zum Datenschutz

Wir nehmen den Schutz Ihrer persönlichen Daten sehr ernst und folgen dabei den aktuellen Vorschriften der Datenschutzgrundverordnung (DSGVO) und dem Bayerischen Datenschutzgesetz (BayDSG). Alle Angaben, die Sie im Zuge der einzelnen Tests und Befragungen machen, sind anonym und werden von uns streng vertraulich behandelt. Sie dienen ausschließlich zu wissenschaftlichen Zwecken und werden nicht an unbefugte Dritte weitergegeben. Die Teilnahme an der Studie ist freiwillig. Sie können Ihre Teilnahme an der Studie jederzeit abbrechen. Ihre Angaben werden automatisch und anonym mit Datum und Uhrzeit gespeichert. Sollten Sie nach Abschluss der Befragung Ihre Einwilligung widerrufen wollen, senden Sie uns bitte eine E-Mail an ipg-gerotest@fau.de, in der Sie das Datum und die Uhrzeit (möglichst genau) Ihrer Teilnahme an der Studie angeben, damit wir Ihre Daten entsprechend löschen können.

Einwilligung

(consent) Ich habe die Information zum Datenschutz gelesen und willige ein, an der Studie teilzunehmen. (1 Zustimmung)

Idcode

Bevor die Studie beginnt, möchten wir Sie bitten, ein persönliches Codewort zu erstellen. Dieses setzt sich aus Ihren Antworten auf die untenstehenden Fragen zusammen. Der Code wird dafür benutzt, die aktuelle Befragung zu anonymisieren, so dass Sie die Möglichkeit haben, Ihre Einwilligung zur Studie jederzeit zurückzuziehen

Es werden Umlaute (ä,ö und ü) und sowohl Groß- als auch Kleinschreibung akzeptiert. Beim letzten Item geben Sie bitte zwei Ziffern ein.

(idcode_1) dritte Buchstabe Ihres Vornamens (z.B. Peter) (idcode_2) erster Buchstabe des Vornames Ihrer Mutter (z.B. Maria)
(idcode_3) dritter Buchstabe des Vornamens Ihrer Mutter (z.B. Maria) (idcode_4) erster Buchstabe des Geburtsnamens Ihrer Mutter (z.B. Meier) (idcode_5) dritter Buchstabe des Geburtsnamens Ihrer Mutter (z.B. Meier) (idcode_6) erster Buchstabe Ihrer Geburtsstadt (z.B. Berlin)
(idcode_7) dritter Buchstabe Ihrer Geburtsstadt (z.B. Berlin)
(idcode_8) der Tag, an dem Sie geboren wurden? (z. B. 07, wenn Sie am 7. August Geburtstag haben)

Soziodemographische Faktoren

Fragen zur eigenen Person

Vielen Dank, dass Sie an der Studie teilnehmen und sich für die Beantwortung unseres Fragebogens Zeit nehmen. Im Folgenden möchten wir Sie nun zunächst bitten, einige Angaben zu Ihrer Person und zu Ihrer Lebenssituation zu beantworten. Diese Angaben sind für uns wichtig, um die unterschiedlichen Meinungen zu künstlicher Intelligenz besser verstehen zu können.

Alter und Geschlecht

Kurzbeschreibung:

Geburtsjahr, biologisches Geschlecht sowie Geschlechtsidentifikation.

Items:

(byear) In welchem Jahr wurden Sie geboren? Bitte geben Sie das vierstellige Geburtsjahr an (z.B. 1970).
_____ (numerisch, vierstellig)

(sex) Welches Geschlecht wurde bei Ihrer Geburt in Ihre Geburtsurkunde eingetragen?

- 0 weiblich
- 1 männlich

(gender) Welchem Geschlecht ordnen Sie sich selbst zu?

- 0 Weiblich
- 1 Männlich
- 2 Transgeschlechtlich
- 3 Anders Geschlecht und zwar: _____ (gender_txt)
- 4 Keine Angabe

Kommentar:

Soziodemographische Daten werden analog zu verschiedenen bevölkerungsrepräsentativen Stichproben (SHARE, DEAS, SOEP) erfasst. Die Geschlechtskategorie „divers“ wird nicht erfasst, da dieser Eintrag in das Geburtenregister erst ab 2017 möglich ist. Umcodieren: Soci alle Werte -1.

Quellen:

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., . . . Zuber, S. (2013). Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *International Journal of Epidemiology*, 42(4), 992–1001. doi:10.1093/ije/dyt088

Kantar Public. 2020. SOEP-Core – 2019: Personenfragebogen, Stichprobe Q (LGB). SOEP Survey Papers 929:Series A. Berlin: DIW/SOEP

Höchster Bildungsabschluss

Kurzbeschreibung:

Höchster Bildungsabschluss

Items:

(educ) Bitte geben Sie an, welchen höchsten Schulabschluss Sie haben. Bitte denken Sie auch an den zweiten Bildungsweg

- 1 Noch Schüler
- 2 Schule beendet ohne Abschluss
- 3 Volks- oder Hauptschulabschluss
- 4 Realschulabschluss (Mittlere Reife, Lyzeum)
- 5 Abschluss Polytechnische Oberschule 10. Klasse (vor 1965: 8. Klasse)
- 6 Fachhochschulreife (Abschluss einer Fachoberschule)
- 7 Abitur, allgemeine oder fachgebundene Hochschulreife (Gymnasium bzw. EOS)
- 8 Hochschulabschluss
- 9 Anderen Abschluss (offene Antwort)

Kommentar:

Soziodemographische Daten werden analog zu verschiedenen bevölkerungsrepräsentativen Stichproben (SHARE, DEAS, SOEP) erfasst. Mögliche Kodierung kann nach ISCED 2011 erfolgen.

Quellen:

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., . . . Zuber, S. (2013). Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *International Journal of Epidemiology*, 42(4), 992–1001. doi:10.1093/ije/dyt088

Erwerbstätigkeit

Kurzbeschreibung:

Erhebung der Erwerbssituation

Items:

(work) Üben Sie derzeit eine Erwerbstätigkeit aus?

Sind Sie derzeit...

- 1 Voll erwerbstätig
- 2 In Teilzeitbeschäftigung
- 3 In Kurzarbeit
- 4 In betrieblicher Ausbildung/Lehre oder betrieblicher Umschulung
- 5 Geringfügig oder unregelmäßig erwerbstätig
- 6 In Altersteilzeit
- 7 Im Freiwilligen Sozialen Jahre/Ökologischen Jahr, im Bundesfreiwilligendienst
- 8 In Rente oder Pension mit geringfügiger Beschäftigung
- 9 In Rente oder Pension
- 10 Nicht erwerbstätig

Kommentar:

Soziodemographische Daten werden analog zu verschiedenen bevölkerungsrepräsentativen erfasst; hier die Erhebung aus dem SOEP-Corona-Fragebogen 2020.

Quellen:

Kühne, S., Kroh, M., Liebig, S. & S. Zinn, (2020): The Need for Household Panel Surveys in Times of Crisis: The Case of SOEP-CoV. In: *Survey Research Methods* 14(2): 195-203.

SOEP-CORE:SOEP-CORE-2020-CORONA Fragebogen Welle 1: <https://www.soep-cov.de/Methodik/>

Haushaltsgröße

Kurzbeschreibung:

Erfassung der Haushaltsgröße

Items:

(hsize) Wie viele Personen leben zur Zeit in Ihrem Haushalt?

- 1 Ich lebe alleine.
- 2 Ich lebe mit einer weiteren Person zusammen.
- 3 Ich lebe mit zwei oder drei Personen zusammen.
- 4 Ich lebe mit vier oder mehr Personen zusammen.

Kommentar:

Erfassung der Haushaltsgröße analog zur VW-Studie.

Quellen:

Lang, F. R., Lessenich, S., & Rothermund, K. (2022). *Altern als Zukunft—eine Studie der Volkswagenstiftung* (p. 173). Springer Nature.

Räumlicher Kontext

Gemeindegröße

Kurzbeschreibung:

Erfassung der Stadt- und Gemeindetypen.

Items:

(area) Was trifft am ehesten auf die Wohngegend zu, in der Sie leben?

- 1 Ich lebe in einer Metropole (über 500.000 EinwohnerInnen)
- 2 Ich lebe in einer Großstadt (über 100.000 EinwohnerInnen)
- 3 Ich lebe am Rand oder in einem Vorort einer Großstadt
- 4 Ich lebe in einer mittelgroßen Stadt (40.000 – 100.000 EinwohnerInnen)
- 5 Ich lebe in einer Kleinstadt (5.000 – unter 40.000 EinwohnerInnen)
- 6 Ich lebe in einem Dorf
- 7 Ich lebe in einem alleinstehenden Haus oder Hof am Land
- 8 keine Angabe

Kommentar:

Stadt- und Gemeindeklassifikation in Anlehnung an Gemeindereferenz des Bundesinstituts für Bau-, Stadt- und Raumforschung (BBSR).

Quellen:

www.bbsr.bund.de

Erreichbarkeit Versorgungsangebote

Kurzbeschreibung:

Erreichbarkeit medizinischer Versorgung und Nahversorgung.

Items:

(preq) Wenn Sie an Ihren Wohnort denken, wie einfach ist es für Sie, die folgenden Versorgungsangebote zu erreichen?

Wie einfach ist es für Sie, ...

(preq_1) zur nächsten Bank oder zum nächsten Geldautomaten zu kommen? (preq_2) zum nächsten Lebensmittelgeschäft oder Supermarkt zu kommen? (preq_3) zu Ihrem Hausarzt oder Ihrer Hausärztin zu kommen?

(preq_4) zur nächsten Apotheke zu kommen? (preq_5) zum nächsten Krankenhaus zu kommen?

Skalierung:

- 1 sehr schwierig
- 2 schwierig
- 3 einfach
- 4 sehr einfach

Kommentar:

Items aus dem Multidimensional Perception of Residential Quality and Neighbourhood Attachment Questionnaire.

Quellen:

Bonaiuto, M., Aiello, A., Perugini, M., Bonnes, M., & Ercolani, A. P. (1999). Multidimensional perception of residential environment quality and neighbourhood attachment in the urban environment. *Journal of Environmental Psychology*, 19, 331-352.

Postleitzahl

Kurzbeschreibung:

Erfassung der Postleitzahl.

Items:

(plz) Bitte geben Sie Ihre Postleitzahl an.

Meine Postleitzahl ist: _____ (numerisch, fünfstellig)

Kommentar:

Wichtig für Differenzierung ländlicher/urbaner Raum sowie Nachrekrutierung

Quellen:

Video – Einführung

Was ist eigentlich Künstliche Intelligenz?

Im Folgenden möchten wir Sie nun bitten, sich das unten eingefügte Video anzuschauen. Das Video soll Ihnen helfen, einen kurzen generellen Einblick in das Thema der künstlichen Intelligenz zu erhalten.

Das Video ist vertont, bitte schalten Sie also die Tonwiedergabe Ihres Computers oder Ihres mobilen Endgeräts ein. Sie können das Video jederzeit pausieren oder auch wiederholt anschauen. Wenn Sie das Video angeschaut und verstanden haben, dann fahren Sie bitte mit der Befragung fort.

[KI-BA_KI-Einfuehrung_mit_Sound.mp4]

KI-Akzeptanz in der Gesundheitsversorgung

Wie ist Ihre Meinung zu Künstlicher Intelligenz?

Vielen Dank, dass Sie sich das Video angeschaut haben. Nun interessieren wir uns dafür, welche Meinung Sie zum Einsatz Künstlicher Intelligenz in der Gesundheitsversorgung haben.

Kurzbeschreibung:

Erfassung der KI-Akzeptanz in der Gesundheitsversorgung.

Items:

Bitte beurteilen Sie die folgenden Aussagen zum Einsatz künstlicher Intelligenz in der Gesundheitsversorgung.

Nr.	Künstliche Intelligenz in der Gesundheitsversorgung ..	Pol.	Dimension
accept1_1	erhöht mein Vertrauen in medizinische Empfehlungen.	+	Vertrauen
accept1_2	verringert mein Vertrauen in medizinische Empfehlungen.	-	Vertrauen
accept1_3	finde ich zuverlässig.	+	Vertrauen
accept1_4	verringert mein Vertrauen in medizinische Diagnosen.	-	Vertrauen
accept1_5	finde ich unsicher.	-	Vertrauen
accept1_6	erfüllt mich mit Angst.	+	Emotion
accept1_7	bereitet mir Unbehagen.	+	Emotion
accept1_8	fühlt sich für mich bedrohlich an.	+	Emotion
accept1_9	ist beklemmend für mich.	+	Emotion
accept1_10	lässt mich sorgenvoll in die Zukunft blicken.	+	Emotion
	Durch den Einsatz von künstlicher Intelligenz in der Gesundheitsversorgung ..		
accept2_11	unterliegen medizinische Entscheidungen meiner Kontrolle.	+	Kontrolle
accept2_12	kann ich medizinische Entscheidungen hinterfragen.	+	Kontrolle
accept2_13	trifft der Mensch die wichtigen medizinischen Entscheidungen.	+	Kontrolle

Anlage 2: AP 3 – Codebuch Akzeptanzstudie Versicherte



accept2_14 unterliegt der Umgang mit medizinischen Empfehlungen meiner Kontrolle.

+ Kontrolle

Anlage 2: AP 3 – Codebuch Akzeptanzstudie Versicherte



accept2_15	habe ich Einfluss auf medizinische Empfehlungen.	+	Kontrolle
accept2_16	profitieren Patienten von maßgeschneiderten Therapien.	+	Nutzen
accept2_17	wird die Gesundheitsversorgung von Patienten verbessert.	+	Nutzen
accept2_18	können Krankheiten besser behandelt werden.	+	Nutzen
accept2_19	profitiert das Gesundheitssystem insgesamt.	+	Nutzen
accept2_20	werden medizinische Empfehlungen für Menschen verbessert.	+	Nutzen

Skalierung:

- 1 trifft überhaupt nicht zu
- 2 trifft eher nicht zu
- 3 weder noch
- 4 trifft eher zu
- 5 trifft genau zu

Kommentar:

Eigene Entwicklung einer KI-Akzeptanzskala in der Versorgung

Quellen:

Hösch et al. (in prep).

Allgemeine KI-Akzeptanz

Kurzbeschreibung:

Erfassung der allgemeinen KI-Akzeptanz.

Items:

Bitte bewerten Sie folgende Aussagen zur künstlichen Intelligenz aus Ihrer Perspektive.

(atai_1) Ich habe Angst vor künstlicher Intelligenz. (atai_2)

Ich vertraue künstlicher Intelligenz.

(atai_3) Künstliche Intelligenz wird die Menschheit zerstören.

(atai_4) Künstliche Intelligenz wird eine Bereicherung für die Menschheit sein. (atai_5)

Künstliche Intelligenz wird für viel Arbeitslosigkeit sorgen.

Skalierung:

0 starke Ablehnung

10 starke Zustimmung

Kommentar:

Erfassung der allgemeinen KI-Akzeptanz durch die ATAI-Skala von Skindermann et al. (2021); originale Skalierung geht von 0-10, hier aufgrund von SoSci allerdings 1-11. Umcodieren: SoSci alle Werte -1.

Quellen:

Sindermann, C., Sha, P., Zhou, M., Wernicke, J., Schmitt, H. S., Li, M., Sariyska, R., Stavrou, M., Becker, B. & Montag, C., (2021). Assessing the Attitude Towards Artificial Intelligence: Introduction of a Short Measure in German, Chinese, and English Language. KI - Künstliche Intelligenz: Vol. 35, No. 1. Springer. (S. 109-118). DOI: 10.1007/s13218-020-00689-0

Video – Stimuli

Künstliche Intelligenz - ein mögliches Anwendungsbeispiel

Wir möchten Sie nun erneut darum bitten, sich ein kurzes Video anzuschauen. Dieses mal geht es um ein mögliches Anwendungsbeispiel einer konkreten KI-basierten Lösung in der Gesundheitsversorgung.

Das Video ist vertont, bitte schalten Sie also die Tonwiedergabe Ihres Computers oder ihres mobilen Endgeräts ein. Sie können das Video jederzeit pausieren oder auch wiederholt anschauen. Wenn Sie das Video angeschaut und verstanden haben, dann fahren Sie bitte mit der Befragung fort.

[Randomisierte Stimuli:

- 01 = EKG-KI-BA_KI-K_mit_Sound.mp4
- 02 = EKG-KI-BA_KI-L_mit_Sound.mp4
- 03 = EKG-KI-BA_ME-K_mit_Sound.mp4
- 04 = EKG-KI-BA_ME-L_mit_Sound.mp4
- 05 = Insulin-KI-BA_KI-K_mit_Sound.mp4
- 06 = Insulin-KI-BA_KI-L_mit_Sound.mp4
- 07 = Insulin-KI-BA_ME-K_mit_Sound.mp4
- 08 = Insulin-KI-BA_ME-L_mit_Sound.mp4
- 09 = Darm-KI-BA_KI-K_mit_Sound.mp4
- 10 = Darm-KI-BA_KI-K_mit_Sound.mp4
- 11 = Darm-KI-BA_ME-K_mit_Sound.mp4
- 12 = Darm-KI-BA_ME-L_mit_Sound.mp4]

Manipulationcheck

Was ist Ihre Meinung zum gezeigten Anwendungsbeispiel?

Vielen Dank, dass Sie sich das Video angeschaut haben. Nun interessieren wir uns dafür, wie Sie das gezeigte KI-basierte Anwendungsbeispiel bewerten und welche Meinung Sie dazu haben. Weiterhin ist es für uns wichtig zu erfahren, ob technische Probleme beim Abspielen des Videos aufgetreten sind und ob die Inhalte verständlich waren.

Kurzbeschreibung: Technische Evaluation

Allgemeine technische Abfrage.

Items:

(tecinfo_1) Konnten Sie das Video abspielen?

(tecinfo_2) Hat der Ton beim Abspielen des Videos funktioniert?

Skalierung:

0 nein

1 ja

Kurzbeschreibung: Manipulationscheck

Manipulationscheck zur Absicherung der experimentellen Bedingungen.

Items:

Bitte denken Sie nun an die KI-Anwendung, die wir Ihnen im vorherigen Video präsentiert haben. Welche der folgenden Aussagen treffen zu?

(manip1_1) Die künstliche Intelligenz führt automatisiert die weitere Behandlung durch.

(manip1_2) Der Arzt oder Patient kann eine eigene vom Systemvorschlag abweichende Behandlung durchführen.

(manip1_3) Die künstliche Intelligenz vergleicht die Daten des Patienten mit den Daten anderer Patienten.

(manip1_4) Die künstliche Intelligenz greift im Hintergrund auf eine große klinische Datenbank zurück.

Skalierung:

0 nein

1 ja

2 weiß ich nicht

(manip2) Was war der medizinische Anwendungskontext der gezeigten KI-Anwendung?

1 Parkinson

2 Diabetes

3 Darmspiegelung (Koloskopie)

4 Zahnwurzelbehandlung

5 Elektrokardiogramm (EKG)

6 Keine der aufgezählten

7 Das weiss ich nicht

Kommentar:

SoSci Werte -1.

Technikakzeptanz

Kurzbeschreibung:

Bewertung der Stimuli mit Blick auf klassische Indikatoren der Technikakzeptanzforschung.

Items:

Bitte geben Sie an, wie Sie die im Video gezeigte KI-Anwendung aus Ihrer Sicht als Patientin oder Patient bewerten.

(tam_1) Ich kann mir vorstellen, diese KI-Anwendung zu nutzen oder an mir anwenden zu lassen. (tam_2) Ich finde diese KI-Anwendung nützlich.

(tam_3) Diese KI-Anwendung würde vieles komfortabler machen. (tam_4) Diese KI-Anwendung ist leicht verständlich.

Skalierung:

- 1 Trifft gar nicht zu
- 2 Trifft eher nicht zu
- 3 Trifft eher zu
- 4 Trifft genau zu

Kommentar:

In Anlehnung an das Technikakzeptanzmodell; abhängige Variablen. Ursprüngliche Skalierung von Davis et al. (1989): 7-stufige Skala (likely-unlikely and anchor points extremely, quite, slightly and neither)

Quellen:

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35, 982–1003.

User Experience

Kurzbeschreibung:

Bewertung der Stimuli mit Blick auf die UX.

Items:

Bewerten Sie bitte die gezeigte KI-Anwendung anhand der folgenden Gegensatzpaare von Eigenschaften.

(ueqs_1) Behindernd <-> unterstützend
(ueqs_2) Kompliziert <-> einfach (ueqs_3)
Ineffizient <-> effizient (ueqs_4) Verwirred
<-> übersichtlich (ueqs_5) Langweilig <->
spannend (ueqs_6) Uninteressant <->
interessant (ueqs_7) Konventionell <->
originell (ueqs_8) Herkömmlich <->
neuartig

Skalierung:

7-stufiges semantisches Differenzial

Kommentar:

Erfassung der UX anhand der pragmatischen und hedonischen Qualität aus der Kurzversion des User Experience Questionnaires (UEQ); abhängige Variable. Die Langversion enthält 26 Items.

Quellen:

Schrepp, Martin; Hinderks, Andreas; Thomaschewski, Jörg (2017): Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). In: IJIMAI 4 (6), pp. 103–108

Persönlichkeit

Herzlichen Dank für die Bewertung der Videos!

Nun haben wir noch einige Fragen zu verschiedenen Aspekten Ihrer Person, die uns dabei helfen, die Haltung gegenüber künstlicher Intelligenz besser zu verstehen. Dabei geht es um verschiedene Bereiche wie Persönlichkeit, Gesundheit oder auch Techniknutzung.

Wir beginnen mit Fragen zur individuellen Persönlichkeit.

Persönlichkeitsdispositionen

Kurzbeschreibung:

Kurzerfassung der Persönlichkeitsdispositionen aus dem Big Five Inventory.

Items:

Inwieweit treffen die folgenden Aussagen auf Sie zu?

Nr.	Ich bin jemand, der"	Polung	Dimension
bfi_1	gründlich arbeitet.	+	Gewissenhaftigkeit
bfi_2	kommunikativ, gesprächig ist.	+	Extraversion
bfi_3	manchmal etwas grob zu anderen ist.	-	Verträglichkeit
bfi_4	originell ist, neue Ideen einbringt.	+	Offenheit
bfi_5	sich oft Sorgen macht.	+	Neurotizismus
bfi_6	zurückhaltend ist.	-	Extraversion
bfi_7	verzeihen kann.	+	Verträglichkeit
bfi_8	eher faul ist.	-	Gewissenhaftigkeit
bfi_9	aus sich herausgehen kann, gesellig ist.	+	Extraversion
bfi_10	künstlerische Erfahrungen schätzt.	+	Offenheit
bfi_11	leicht nervös wird.	+	Neurotizismus
bfi_12	Aufgaben wirksam und effizient erledigt.	+	Gewissenhaftigkeit
bfi_13	rücksichtsvoll und freundlich mit anderen umgeht.	+	Verträglichkeit
bfi_14	eine lebhaftere Phantasie, Vorstellungen hat.	+	Offenheit
bfi_15	entspannt ist, mit Stress gut umgehen kann.	-	Neurotizismus

Skalierung:

1 trifft überhaupt nicht zu 7
trifft voll und ganz zu

Kommentar:

Das Instrument wurde basierend auf dem Big Five Inventar im Rahmen des SOEP entwickelt und eingesetzt.

Quellen:

Lang, F. R., Hahne, D., Gymbel, S., Schröpper, S., & Lutsch, K. (2005). Erfassung des kognitiven Leistungspotenzials und der "Big Five" mit Computer-Assisted-Personal-Interviewing (CAPI): zur Reliabilität und Validität zweier ultrakurzer Tests und des BFI-S. *DIW Research Notes*, 9, 2005.

Zeitperspektive

Kurzbeschreibung:

Ultrakurze Erfassung der wahrgenommenen verbleibenden Spanne an Lebenszeit.

Items:

Inwiefern treffen die folgenden Aussagen bezüglich Zeit und Zukunft auf Sie zu?

(ftps_1) Auf mich warten viele Möglichkeiten in der Zukunft. (ftps_2)

Meine Zukunft ist voller Möglichkeiten.

(ftps_3) Mit zunehmendem Alter beginne ich, die Zeit als begrenzt zu erleben.

Skalierung:

1 trifft ganz und gar nicht zu

7 trifft voll und ganz zu

Kommentar:

Die Skala mit 10 Items wurde von Rohr et al. (2017) entwickelt, die Auswahl der Items für die ultrakurze Erfassung basiert auf Gellert et al. (2012). Die 7-stufige Skalierung basiert auf der VW Studie.

Quellen:

Lang, F. R., & Carstensen, L. L. (2002). Time counts: future time perspective, goals, and social relationships. *Psychology and aging, 17*(1), 125.

Rohr, M. K., John, D. T., Fung, H. H., & Lang, F. R. (2017). A three-component model of future time perspective across adulthood. *Psychology and Aging, 32*(7), 597.

Gellert, P., Ziegelmann, J. P., Lippke, S., & Schwarzer, R. (2012). Future time perspective and health behaviors: Temporal framing of self-regulatory processes in physical exercise and dietary behaviors. *Annals of Behavioral Medicine, 43*(2), 208-218.

Obsoleszenzerleben

Kurzbeschreibung:

Erfassung der wahrgenommenen Obsoleszenz.

Items:

Alle Menschen werden älter. Was Älterwerden für den Einzelnen jedoch bedeutet, kann sehr unterschiedlich sein. Wie erleben Sie Veränderungen, die mit dem Älterwerden einhergehen können?

(obs_1) Ich habe zunehmend das Gefühl, den Anschluss an die heutige Zeit verpasst zu haben. (obs_2) Ich bin voll auf der Höhe der Zeit.

(obs_3) Ich komme mit der heutigen Lebensweise immer schlechter zurecht.

(obs_4) Für die Auffassungen der jüngeren Generation habe ich immer weniger Verständnis. (obs_5) Das Leben wird für mich immer komplizierter und schwerer zu durchschauen.

Skalierung:

- 1 trifft überhaupt nicht zu
- 2 trifft eher nicht zu
- 3 weder noch
- 4 trifft eher zu
- 5 trifft voll und ganz zu

Kommentar:

Die Items wurden in Anlehnung an Brandstädter & Wentura (1994) entwickelt und werden zusammen mit den Fragen zum subjektiven Alterserleben erhoben.

Quellen:

Brandstädter, J., & Wentura, D. (1994). Veränderungen der Zeit- und Zukunftsperspektive im Übergang zum höheren Erwachsenenalter: entwicklungspsychologische und differentielle Aspekte. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 26(1), 2–21.

Subjektives Alterserleben

Kurzbeschreibung:

Erfassung des subjektiven Alterserlebens.

Items:

Alle Menschen werden älter. Was Älterwerden für den Einzelnen jedoch bedeutet, kann sehr unterschiedlich sein. Wie erleben Sie Veränderungen, die mit dem Älterwerden einhergehen können?

(pgcms_1) Je älter ich werde, desto schlimmer wird alles. (pgcms_2) Ich habe noch genauso viel Schwung wie letztes Jahr. (pgcms_3) Je älter ich werde, desto weniger nützlich bin ich.

(pgcms_4) Mit zunehmendem Alter ist mein Leben besser, als ich erwartet habe. (pgcms_5) Ich bin jetzt genauso glücklich, wie ich es in jungen Jahren war.

Skalierung:

- 1 Trifft gar nicht zu
- 2 Trifft eher nicht zu
- 3 Trifft eher zu
- 4 Trifft genau zu

Kommentar:

Die Items stammen aus der PGCMS (Lawton, 1975) und werden in verschiedenen Panelstudien (z.B. HRS, DEAS) zur Erfassung des subjektiven Alterserlebens und des persönlichen Altersbildes genutzt. Die Items werden zusammen mit den Fragen zur Obsoleszenz erhoben.

Quellen:

Lawton, M.P. (1975). The Philadelphia Geriatric Center Morale Scale: A revision. *Journal of Gerontology*, 30, 85-89.

Novelty

Kurzbeschreibung:

Erfassung der intellektuellen Neugier und Wissberigkeit (curiosity).

Items:

Im Folgenden möchten wir gerne erfahren, wie Sie mit neuen Dingen und Erfahrungen umgehen. Bitte schätzen Sie für jede der folgenden Aussagen ein, wie gut diese auf Sie zutrifft.

(nov_1) Ich profitiere davon, Neues zu lernen.

(nov_2) Bei neuen Entwicklungen auf dem Laufenden zu bleiben, unterstützt einen unabhängigen Lebensstil im Alter.

(nov_3) Sich mit neuen Dingen zu beschäftigen, verbessert meine Ressourcen (z.B. Wissen, soziale Kontakte, ...)

(nov_4) Andere werden mich verurteilen, wenn ich mit neuen Technologien nicht zurecht komme. (nov_5)

Es kostet zu viel Zeit und Energie, etwas Neues benutzen zu lernen.

(nov_6) Sich auf neue Dinge einzulassen, stört meine täglichen Routinen. (nov_7) Ich probiere gerne neue Dinge aus.

(nov_8) Ich verlasse gerne meine Komfortzone.

(nov_9) Ich bevorzuge bekannte Aktivitäten gegenüber neuen Abenteuern.

Skalierung:

- 1 trifft überhaupt nicht zu
- 2 trifft eher nicht zu
- 3 weder noch
- 4 trifft eher zu
- 5 trifft voll und ganz zu

Kommentar:

Eigene Items zu novelty

Quellen:

Neugier

Kurzbeschreibung:

Erfassung der intellektuellen Neugier und Wissberigkeit (curiosity).

Items:

Wie ist Ihre persönliche Erfahrung mit Blick auf neue Situationen und Erfahrungen?

(cur_1) Ich sehe herausfordernde Situationen als eine Gelegenheit, zu lernen und mich weiterzuentwickeln.

(cur_2) Ich bin stets auf der Suche nach Erlebnissen, die die Art und Weise, wie ich über mich selbst und die Welt denke, herausfordern.

(cur_3) Ich suche Situationen auf, in denen es wahrscheinlich ist, dass ich sehr gründlich über etwas nachdenken muss.

(cur_4) Ich genieße es, etwas über Themen zu lernen, die mir nicht vertraut sind. (cur_5) Ich finde es faszinierend, neue Informationen zu lernen.

Skalierung:

1 trifft ganz und gar nicht zu
7 trifft voll und ganz zu

Kommentar:

Die Items stammen aus der curiosity Skala von Kashdan et al (2018) und wurden in der VW Studie (7- stufig) eingesetzt.

Quellen:

Kashdan, T. B., Stikma, M. C., Disabato, D. J., McKnight, P. E., Bekier, J., Kaji, J., & Lazarus, R. (2018). The Five-Dimensional Curiosity Scale: Capturing the bandwidth of curiosity and identifying four unique subgroups of curious people. *Journal of Research in Personality*, 73, 130–149.

Lebenszufriedenheit

Kurzbeschreibung:

Einschätzung der eigenen subjektiven Lebenszufriedenheit.

Items:

(lifesat) Wie zufrieden sind Sie, alles in allem, mit Ihrem Leben?

Skalierung:

0 ganz und gar unzufrieden 10
ganz und gar zufrieden

Kommentar:

Erfassung analog zur Erhebung in diversen Panelstudien; Skalierung geht i.d.R. von 0-10, wird aufgrund von SoSci entsprechend mit 1-11 skaliert. Umkodieren: SoSci alle Werte -1

Quellen:

Engstler, H., Alcantara, A. L., Luitjens, M., & Klaus, D. (2019). German Ageing Survey, Deutscher Alterssurvey (DEAS): Documentation of instruments and variables 1996 – 2017. doi: 10.5156/DEAS.1996-2017.D.002

Gesundheitliche Ressourcen

Gesundheit

Vielen Dank für die Beantwortung der Fragen zur Persönlichkeit!

Wir glauben, dass die Haltung gegenüber künstlicher Intelligenz in der Gesundheitsversorgung auch mit gesundheitlichen Aspekten zusammenhängt. Mit den nächsten Fragen möchten wir mehr darüber erfahren.

Subjektive Gesundheit

Kurzbeschreibung:

Einschätzung der eigenen subjektiven Gesundheit

Items:

(health) Wie würden Sie Ihren gegenwärtigen Gesundheitszustand beschreiben?

- 1 schlecht
- 2 weniger gut
- 3 zufriedenstellend
- 4 gut
- 5 sehr gut

Kommentar:

Erfassung der subjektiven Gesundheit analog zur Erhebung im SOEP.

Quellen:

SOEP-IS Group, 2021. SOEP-IS 2019 – Questionnaire for the SOEP Innovation Sample. SOEP Survey Papers 956: Series A – Survey Instruments (Erhebungsinstrumente). Berlin: DIW Berlin/SOEP

COVID-19 Infektion

Kurzbeschreibung:

Erfassung einer zurückliegenden COVID-19 Infektion.

Items:

(covid) Waren Sie in den letzten 6 Monaten mit COVID-19 infiziert?

- 1 Nein, ich war nicht wesentlich infiziert.
- 2 Ja, ich hatte einen symptomfreien Verlauf.
- 3 Ja, ich hatte einen milden Verlauf.
- 4 Ja, ich hatte einen schweren Verlauf.

Kommentar:

Analog zur Erfassung in der IPG-Corona-Studie sowie diversen Paneldaten.

Quellen:

Stuth, S. (2022). Deutscher Alterssurvey (DEAS): Kurzbeschreibung des SUF DEAS 2020/21, Version 1.1. Berlin: Deutsches Zentrum für Altersfragen. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-79742-5>

Funktionale Beeinträchtigungen

Kurzbeschreibung:

Globaler Indikator zur Analyse von gesundheitlichen funktionalen Beeinträchtigungen.

Items:

(gali) In welchem Maße waren Sie während der letzten sechs Monate wegen eines gesundheitlichen Problems bei alltäglichen Verrichtungen eingeschränkt?

- 1 Stark eingeschränkt
- 2 Eingeschränkt, aber nicht stark
- 3 Nicht eingeschränkt

Skalierung:

1 = „zutreffend“

Kommentar:

Erfassung von gesundheitlichen Einschränkungen (GALI) analog zur Erhebung im SHARE.

Quellen:

- Verbrugge, L. M. (1997). A global disability indicator. *Journal of aging studies*, 11(4), 337-362. doi: 10.1016/S0890-4065(97)90026-8
- Robine, J. M., Jagger, C., & Euro-REVES Group. (2003). Creating a coherent set of indicators to monitor health across Europe: the Euro-REVES 2 project. *European journal of public health*, 13(suppl_3), 6-14. doi: 10.1093/eurpub/13.suppl_3.6
- Van Oyen, H., Van der Heyden, J., Perenboom, R., & Jagger, C. (2006). Monitoring population disability: evaluation of a new Global Activity Limitation Indicator (GALI). *Sozial-und Präventivmedizin*, 51(3), 153-161. doi: 10.1007/s00038-006-0035-y

Einsamkeit

Kurzbeschreibung:

Ultrakurze Erfassung von Einsamkeitserleben.

Items:

Bitte geben Sie für jede Frage an, wie oft Sie sich so fühlen. (**lone_1**) Wie häufig vermissen Sie Gesellschaft?

(**lone_2**) Wie häufig fühlen Sie sich isoliert von anderen? (**lone_3**) Wie häufig fühlen Sie sich einsam?

Skalierung:

- 1 niemals
- 2 selten
- 3 manchmal
- 4 häufig

Kommentar:

Basiert auf Kurzversion der R-UCLA, eingesetzt in SHARE.

Quellen:

Hughes, M. E., Waite, L. J., Hawkey, L. C., & Cacioppo, J. T. (2004). A short scale for measuring loneliness in large surveys: Results from two population-based studies. *Research on aging, 26*(6), 655- 672.

Ängstlichkeit und Depression

Kurzbeschreibung:

Ultrakurze Erfassung depressiver Symptome und generalisierter Ängstlichkeit.

Items:

Jetzt geht es um die letzten 2 Wochen. Wie oft fühlten Sie sich da durch die folgenden Beschwerden beeinträchtigt?

(**anx_1**) Wenig Interesse oder Freude an Ihren Tätigkeiten

(**anx_2**) Niedergeschlagenheit, Schwermut oder Hoffnungslosigkeit (**anx_3**)
Nervosität, Ängstlichkeit oder Anspannung

(**anx_4**) Nicht in der Lage sein, Sorgen zu stoppen oder zu kontrollieren

Skalierung:

- 1 überhaupt nicht
- 2 an mehreren Tagen
- 3 mehr als die Hälfte der Tage
- 4 beinahe jeden Tag

Kommentar:

Ultrakurzform des Gesundheitsfragebogens für Patienten (PHQ-4) zur Erfassung der diagnostischen Kernkriterien für depressive Störungen und für die Generalisierte Angststörung nach DSM-IV.

Quellen:

B. Löwe, I. Wahl, M. Rose, C. Spitzer, H. Glaesmer, K. Wingenfeld, A. Schneider, E. Brähler: A 4- item measure of depression and anxiety: Validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. In: *Journal of Affective Disorders*. 122, 2010, S. 86- 95.

Diagnosen

Kurzbeschreibung:

Erfassung von diagnostizierten Krankheiten.

Items:

Hat Ihnen ein Arzt schon einmal gesagt, dass Sie unter einer der dort aufgeführten Krankheiten leiden?

- (disease_1) Erhöhte Blutfettwerte (Cholesterinwerte)
- (disease_22) hohe Blutzuckerwerte
- (disease_23) Diabetes Typ 1
- (disease_24) Diabetes Typ 2
- (disease_3) Bluthochdruck
- (disease_4) Herzinfarkt, Angina pectoris
- (disease_5) Herzschwäche (Herzinsuffizienz), einschließlich Durchblutungsstörungen am Herzen (disease_6) Schlaganfall
- (disease_7) Durchblutungsstörungen im Gehirn
- (disease_8) Durchblutungsstörungen in den Beinen
- (disease_9) Gelenkverschleiß (Arthrose) der Hüft- oder Kniegelenke bzw. der Wirbelsäule (disease_10) Osteoporose
- (disease_11) Entzündliche Gelenk- oder Wirbelsäulenerkrankung (Arthritis oder Rheuma) (disease_12) Chronische Lungenerkrankung (z.B. chronische Bronchitis, Lungenemphysem) (disease_13) Krebserkrankung, bösartiger Tumor (einschließlich Leukämie)
- (disease_14) Magengeschwür, Zwölffingerdarmgeschwür (disease_15) Inkontinenz
- (disease_16) Seelische Erkrankung (z.B. Angstzustände, Depression, Psychose) (disease_17) Parkinson Erkrankung
- (disease_18) Glaukom (grüner Star) oder Makuladegeneration
- (disease_19) Andere chronische Erkrankung oder Gesundheitsstörung (gemeint sind nur länger andauernde oder wiederkehrende Erkrankungen)
- (disease_20) und zwar: _____ (disease_txt)
- (disease_21) Keine Erkrankung

Skalierung:

1 ja

Kommentar:

Erfassung analog zu bevölkerungsrepräsentativen Paneldaten; hier DEAS.

Quellen:

Engstler, H., Alcantara, A. L., Luitjens, M., & Klaus, D. (2019). German Ageing Survey, Deutscher Alterssurvey (DEAS): Documentation of instruments and variables 1996 – 2017. doi: 10.5156/DEAS.1996-2017.D.002

Gesundheitsverhalten

Krankenhausaufenthalte

Kurzbeschreibung:

Erfassung Krankenhausaufenthalte in den vergangenen 12 Monaten.

Items:

(hosp) Wie viele Nächte waren Sie in den vergangenen 12 Monaten zur stationären Behandlung in einem Krankenhaus aufgenommen?

Anzahl der Nächte: _____ (numerisch)

Kommentar:

Objektiver Gesundheitsindikator in Anlehnung an diverse Panelstudien

Quellen:

Bellach, B. M., Knopf, H., & Thefeld, W. (1998). The German Health Survey. 1997/98]. Gesundheitswesen (Bundesverband der Ärzte des Öffentlichen Gesundheitsdienstes (Germany)), 60, S59.

Arztkontakte

Kurzbeschreibung:

Erfassung Inanspruchnahme medizinischer Versorgung.

Items:

Wie häufig haben Sie die folgenden Ärzte in den letzten 12 Monaten in Anspruch genommen? Bitte zählen Sie auch Hausbesuche und Video-/Telefonsprechstunden mit.

(doc_1) praktischer Arzt / Arzt für Allgemeinmedizin

(doc_2) Zahnarzt

(doc_3) Internist

(doc_4) Frauenarzt / Gynäkologe (do_5)

Augenarzt

(doc_6) Orthopäde

(doc_7) Hals-Nasen-Ohren-Arzt

(doc_8) Nervenarzt / Psychiater

(doc_9) Hautarzt

(doc_10) Urologe

(doc_11) Sonstiger Arzt

Skalierung:

1 nie

2 einmal

3 2-3 mal

4 4-6 mal

5 7-12 mal

6 öfter

Kommentar:

Erfassung der Arztkontakte; eingesetzt im DEAS.

Quellen:

Bellach, B. M., Knopf, H., & Thefeld, W. (1998). The German Health Survey. 1997/98]. Gesundheitswesen (Bundesverband der Ärzte des Öffentlichen Gesundheitsdienstes (Germany)), 60, S59.

Vorsorge

Kurzbeschreibung:

Erfassung des gesundheitsbezogenen Vorsorgeverhaltens.

Items:

(prev) Ich nehme regelmäßig an Vorsorge-Untersuchungen (z.B. Gesundheits-Check-Up, Krebsfrüherkennung, Zahnkontrolle) teil.

Skalierung:

- 1 Trifft gar nicht zu
- 2 Trifft eher nicht zu
- 3 Trifft eher zu
- 4 Trifft genau zu

Kommentar:

Gesundheitsbezogenes Vorsorgeverhalten; eingesetzt im DEAS.

Quellen:

Bellach, B. M., Knopf, H., & Thefeld, W. (1998). The German Health Survey. 1997/98]. Gesundheitswesen (Bundesverband der Ärzte des Öffentlichen Gesundheitsdienstes (Germany)), 60, S59.

Krankenversicherung

Kurzbeschreibung:

Erfassung Versicherungsstatus.

Items:

(ins) Wie sind Sie derzeit krankenversichert?

- 1 Ich bin aktuell nicht krankenversichert.
- 2 Gesetzliche Krankenversicherung
- 3 Gesetzliche Krankenversicherung mit privater Zusatzversicherung
- 4 Private Krankenversicherung
- 5 Sonstiges: _____ (ins_txt)

Kommentar:

Eigene Abfrage.

Arzt-Patienten-Beziehung

Kurzbeschreibung:

Erfassung der allgemeinen Arzt-Patienten-Beziehung aus Patientensicht.

Items:

Wie ist Ihre Haltung zu Ärztinnen und Ärzten?

(docrel_1) Manchmal kümmern sich Ärzte mehr darum, was für sie bequem ist, als um die medizinischen Bedürfnisse ihrer Patienten.

(docrel_2) Ärzte sind sehr gründlich und vorsichtig.

(docrel_3) Ärztlichen Entscheidungen vertraue ich voll und ganz.

(docrel_4) Ein Arzt würde mich niemals über irgendetwas täuschen.

(docrel_5) Alles in allem vertraue ich Ärzten voll und ganz.

Skalierung:

- 1 trifft gar nicht zu
- 2 trifft eher nicht zu
- 3 weder noch
- 4 trifft eher zu
- 5 trifft genau zu

Kommentar:

Erfassung der allgemeinen Arzt-Patienten-Beziehung aus Patientensicht; Kurzversion der WFPTS Skala nach Hall et al. 2002.

Quellen:

Dugan, E., Trachtenberg, F., & Hall, M. A. (2005). Development of abbreviated measures to assess patient trust in a physician, a health insurer, and the medical profession. *BMC health services research*, 5, 64. <https://doi.org/10.1186/1472-6963-5-64>

Hall, M. A., Camacho, F., Dugan, E., & Balkrishnan, R. (2002). Trust in the medical profession: conceptual and measurement issues. *Health services research*, 37(5), 1419–1439. <https://doi.org/10.1111/1475-6773.01070>

Soziale Beziehungen

Sozialen Beziehungen

Soziale Kontakte und Beziehungen zu anderen Menschen können vollkommen unterschiedlich aussehen und eine wichtige Rolle spielen, wenn es um die Bewertung von Angeboten in der Gesundheitsversorgung geht. Ihre Angaben zu den folgenden Fragen helfen uns dabei, dies genauer zu verstehen.

Zufriedenheit mit Beziehungen

Kurzbeschreibung:

Erfasst wird die globale Zufriedenheit mit Familienbeziehungen und Freunden.

Items:

(famsat) Wie zufrieden sind Sie mit Ihren Familienbeziehungen? (friendsat)
Wie zufrieden sind Sie mit Ihrem Freundeskreis?

Skalierung:

1 sehr unzufrieden
7 sehr zufrieden

Kommentar:

Items angelehnt an Panelstudien (SOEP, NRW80+, DEAS) sowie an VW Studie.

Quellen:

Lang, F. R., Lessenich, S., & Rothermund, K. (2022). *Altern als Zukunft–eine Studie der Volkswagenstiftung* (p. 173). Springer Nature.

Soziale Unterstützung

Kurzbeschreibung:

Erfasst wird die globale erhaltene Unterstützung von Familienbeziehungen und Freunden.

Items:

(famsupp) Wie häufig erhalten Sie Unterstützung und Hilfe von Ihrer Familie? **(friendsupp)** Wie häufig erhalten Sie Unterstützung und Hilfe aus Ihrem Freundeskreis?

Skalierung:

1 sehr selten

7 sehr häufig

Kommentar:

Items angelehnt an Panelstudien (SOEP, NRW80+, DEAS), in diesen allerdings erfasst im Rahmen von Netzwerkgeneratoren.

Quellen:

Engstler, H., Alcantara, A. L., Luitjens, M., & Klaus, D. (2019). German Ageing Survey, Deutscher Alterssurvey (DEAS): Documentation of instruments and variables 1996 – 2017. doi: 10.5156/DEAS.1996-2017.D.002

Anzahl wichtiger Beziehungen

Kurzbeschreibung:

Erfasst wird die Anzahl bedeutsamer Beziehungspartner.

Items:

(npeople) Wenn Sie nun insgesamt an die Menschen denken, die Ihnen wichtig sind und mit denen Sie in Kontakt stehen – um wieviele Personen handelt es sich dabei?

Anzahl der mir wichtigen Personen: _____ (numerisch)

Kommentar:

Indikator sozialer Ressourcenstrukturen und Eingebundenheit; Frage angelehnt an Panelstudien, häufig Eröffnung für Netzwerkgenerator (SOEP, NRW80+, DEAS).

Quellen:

Malter, F. and Börsch-Supan, A.(Eds.)(2013). SHARE Wave 4: Innovations & Methodology.
Munich: MEA, Max Planck Institute for Social Law and Social Policy.

Digitalisierung und Technik

Umgang mit der Digitalisierung

Nun würden wir von Ihnen noch gerne mehr über Ihre Meinung zu moderner Technik sowie zu Ihrer Internetnutzung erfahren. Bitte beantworten Sie uns die nachfolgenden Fragen.

Internetnutzungsverhalten

Kurzbeschreibung:

Erfassung der Internetnutzung.

Items:

Wie häufig nutzen Sie das Internet für folgende Zwecke?

- (netuse_1) Kontakt mit Freunden, Bekannten und Verwandten (z.B. E-Mail, Facebook, Chat, Videotelefonie)
- (netuse_2) Suche nach sozialen Kontakten (z.B. Freunde, Partner, Gleichgesinnte) (netuse_3) Suche nach Informationen (z.B. Nachrichten, Ratgeber-Seiten, Wikipedia)
- (netuse_4) Bankgeschäfte (z.B. Online-Banking, Banking-Apps, mobiles Bezahlen per Smartphone) (netuse_5) Unterhaltung und Kultur (z.B. Musik hören, Filme schauen, Spiele spielen, Fernsehen) (netuse_6) Einkaufen (z.B. Amazon, Ebay, Online-Apotheke, Lebensmittellieferung)
- (netuse_7) Eigene Inhalte erzeugen (z.B. Texte, Bilder, Musik, Blogs, Webseiten, Online-Verkäufe) (netuse_8) Führen von Beratungsgesprächen mit Ärzt/innen oder Therapeut/innen über eine Online Plattform
- (netuse_9) Verwenden von Anwendungen zur Überwachung und Steigerung von Gesundheit und Wohlbefinden (z.B. Schlafanalyse, Achtsamkeits-Apps, Herzfrequenzmesser, Schrittzähler) (netuse_10) Kontakt mit Behörden und Dienstleistern (z.B. Dokumente beantragen, Terminvereinbarungen)
- (netuse_11) Religiöse Zwecke (z.B. Online-Gottesdienste)

Skalierung:

- 1 nie
- 2 seltener als einmal im Monat
- 3 1-3x im Monat
- 4 1x in der Woche
- 5 Mehrmals in der Woche
- 6 täglich

Kommentar:

Erfassung des Internetnutzungsverhalten analog zur Erhebung im DEAS. Nur für Online-Befragung, bzw. erneute Nachfrage bei CAPI-Befragung, wenn fehlender Internetzugang nicht Motivation für persönliche Befragung ist.

Quellen:

Engstler, H., Alcantara, A. L., Luitjens, M., & Klaus, D. (2019). German Ageing Survey, Deutscher Alterssurvey (DEAS): Documentation of instruments and variables 1996 – 2017. doi: 10.5156/DEAS.1996-2017.D.002

Technikmotivation

Kurzbeschreibung:

Erfassung der persönlichen Technikmotivation.

Items:

Wie ist Ihre Haltung zu moderner Technik?

Nr.		Dimension
stai_1	Ich strenge mich so lange an, bis ein neues Gerät funktioniert, wie ich es will.	TGE
stai_2	Ich vertraue moderner Technik.	PST
stai_3	Die Nutzung moderner Technik hilft mir, wichtige Entscheidungen zu treffen.	PAU
stai_4	Ich übe so lange mit einem neuen Gerät, bis ich dieses optimal benutzen kann.	TGE
stai_5	Technischen Neuerungen sehe ich mit Zuversicht entgegen.	PST
stai_6	Die Nutzung moderner Technik hilft mir bei der Bewältigung des Alltags.	PAU

Anlage 2: AP 3 – Codebuch Akzeptanzstudie Versicherte

- stai_7** Ich verstärke meine Anstrengung, wenn ein neues Gerät schwieriger zu bedienen ist als erwartet.
- stai_8** Ich vertraue darauf, dass neue technische Innovationen hohen Sicherheitsstandards genügen.



TGE
PST

Anlage 2: AP 3 – Codebuch Akzeptanzstudie Versicherte



stai_9 Die Nutzung moderner Technik hilft mir, mein Leben unabhängig zu führen.

PAU

Anlage 2: AP 3 – Codebuch Akzeptanzstudie Versicherte



TGE

stai_10 Wenn ein neues technisches Gerät nicht so funktioniert wie ich es will, spornt mich das zu mehr Anstrengung an.

Anlage 2: AP 3 – Codebuch Akzeptanzstudie Versicherte



stai_11 Moderne Technik gibt mir ein Gefühl der Sicherheit.

PST

Anlage 2: AP 3 – Codebuch Akzeptanzstudie Versicherte



stai_12 Die Nutzung moderner Technik hilft mir, meine täglichen Aufgaben effektiver zu bewältigen.

PAU

Anlage 2: AP 3 – Codebuch Akzeptanzstudie Versicherte



stai_13	Alles in allem nutze ich moderne Technik sehr häufig.	TC
stai_14	Ich interessiere mich für technische Neuerungen.	TC
stai_15	Alles in allem schätze ich mich als kompetent im Umgang mit moderner Technik ein.	TC

Skalierung:

- 1 trifft überhaupt nicht zu
- 2 trifft eher nicht zu
- 3 weder noch
- 4 trifft eher zu
- 5 trifft genau zu

Kommentar:

Erfassung der Technikmotivation über STAI (Kamin & Lang, 2013); die Items 13-15 dienen zu Validierung und können als Outcome genutzt werden.

Quellen:

Kamin, S. T., & Lang, F. R. (2013). The Subjective Technology Adaptivity Inventory (STAI): A motivational measure of technology usage in old age. *Gerontechnology*, 12(1), 16–25.

Digitalkompetenz

Kurzbeschreibung:

Erfassung der Digitalkompetenz.

Items:

Auf Computern und anderen Endgeräten kann man verschiedene Anwendungsmöglichkeiten nutzen. Bitte geben Sie an, inwieweit die Aussagen auf der Liste auf Sie persönlich zutreffen.

(comp_1) Ich kann Internetrecherchen (z. B. bei Google) durchführen.

(comp_2) Ich kann seriöse von unseriösen Nachrichten im Internet unterscheiden.

(comp_3) Ich kann ein Nutzerkonto erstellen, um digitale Online-Dienste (z. B. E-Mails, soziale Medien) zu nutzen.

(comp_4) Ich kann Ordner und Dateien in einer Cloud-Anwendung (z. B. Dropbox, iCloud, Google Drive) via Link mit anderen teilen.

(comp_5) Ich kann eine Videokonferenz über Tools für virtuelle Treffen mit anderen einrichten, z. B. über Microsoft Teams oder Zoom.

(comp_6) Ich bin in der Lage, auf den offiziellen Webseiten von Regierungsstellen etc. passende Informationen und Formulare zu finden (z. B. Bürgeramt, Bundesagentur für Arbeit).

(comp_7) Ich kann digitale Dienstleistungen nutzen (z.B. Dokumente beantragen, Terminvereinbarungen bei Ärzten).

(comp_8) In verschiedenen sozialen Medien habe ich Online-Profile, auf denen ich mich entsprechend darstelle (z. B. berufliche, öffentliche oder private Profile wie Facebook, Xing, usw.).

(comp_9) Ich kann im Internet frei nutzbare Inhalte von urheberrechtlich geschützten Inhalten unterscheiden.

(comp_10) Ich kann verdächtige E-Mails oder Pop-Ups erkennen, die Datenmissbrauch oder Schaden auf meinem digitalen Gerät verursachen können.

Skalierung:

- 1 Trifft gar nicht zu
- 2 Trifft eher nicht zu
- 3 Trifft eher zu
- 4 Trifft genau zu

Kommentar:

Erfassung des Digitalkompetenz; Auswahl der Items aus D21-Digital-Index 2022. Nur für Online- Befragung.

Quellen:

<https://initiated21.de/d21index21-22/>

Abschluss der Befragung

Herzlichen Dank!

Sie haben alle Fragen beantwortet, vielen Dank für Ihre Mühe! Zum Abschluss der Befragung möchten wir gerne sicherstellen, dass wir Ihre Angaben für die wissenschaftliche Auswertung verwenden können.

Reluktanz

Kurzbeschreibung:

Seriousness Checks zur Verbesserung der Datenqualität – Reluktanz beim Ausfüllen.

Items:

(SC02) Jetzt mal Hand aufs Herz: Haben Sie gerne an dieser Studie teilgenommen? Bitte antworten Sie ehrlich - Ihre Antwort hat keine Konsequenzen für Sie!

Skalierung:

- 1 Nein
- 2 Eher Nein
- 3 Eher Ja
- 4 Ja

Kommentar:

Die Items ermöglichen eine Einschätzung der Datenqualität; Anlehnung an Aust et al. (2013); standardmäßig implementiert in SoSci.

Quellen:

Aust, F., Diederhufen, B., Ullrich, S. *et al.* Seriousness checks are useful to improve data validity in online research. *Behav Res* **45**, 527–535 (2013). <https://doi.org/10.3758/s13428-012-0265-2>

Zuverlässigkeit

Kurzbeschreibung:

Seriousness Checks zur Verbesserung der Datenqualität – Zuverlässigkeit der Angaben.

Items:

(SC04) Haben Sie alle Aufgaben so ausgeführt, wie in den jeweiligen Instruktionen gebeten? Bitte antworten Sie auch hier ehrlich - Auch diese Antwort hat für Sie keine Konsequenzen!

Skalierung:

- 1 Ich habe alle Aufgaben, wie in den Instruktionen verlangt, bewältigt.
- 2 __Manchmal__ habe ich irgendetwas geklickt, weil ich unmotiviert war oder mich einfach nicht ausgedenkt habe.
- 3 Ich habe __häufig__ irgendetwas angeklickt, damit ich schnell fertig werde.

Kommentar:

Die Items ermöglichen eine Einschätzung der Datenqualität; Anlehnung an Aust et al. (2013); standardmäßig implementiert in SoSci.

Quellen:

Aust, F., Diederhufen, B., Ullrich, S. *et al.* Seriousness checks are useful to improve data validity in online research. *Behav Res* **45**, 527–535 (2013). <https://doi.org/10.3758/s13428-012-0265-2>

Ablenkung

Kurzbeschreibung:

Seriousness Checks zur Verbesserung der Datenqualität – Ablenkung beim Ausfüllen.

Items:

(SC05) Konnten Sie den Fragebogen in einem ausfüllen, ohne abgelenkt zu werden? Bitte antworten Sie auch hier ehrlich.

Skalierung:

- 1 Ja, ich habe der Studie vollste Aufmerksamkeit geschenkt und sie in einem durch gemacht.
- 2 Ich wurde durch meine Umwelt (Personen, Geräusche, etc.) __einmal__ kurz abgelenkt.
- 3 Ich wurde durch meine Umwelt (Personen, Geräusche, etc.) __mehrmals__ abgelenkt.

Kommentar:

Die Items ermöglichen eine Einschätzung der Datenqualität; Anlehnung an Aust et al. (2013); standardmäßig implementiert in SoSci.

Quellen:

Aust, F., Diedenhofen, B., Ullrich, S. *et al.* Seriousness checks are useful to improve data validity in online research. *Behav Res* **45**, 527–535 (2013). <https://doi.org/10.3758/s13428-012-0265-2>

Verlosung und Kontakt

Kurzbeschreibung:

Anonymisierte Teilnahme an Verlosung und Bereitschaft für weitere Studien kontaktiert zu werden.

Items:

Nochmals vielen Dank für Ihre Teilnahme an unserer Studie!

Sie haben nun die Möglichkeit, an der Verlosung von 100 Gutscheinen im Wert von 20 EUR teilzunehmen. Bitte geben Sie hierzu im nachfolgenden Feld Ihre Email-Adresse an unter der wir Sie im Gewinnfall kontaktieren dürfen. Weiterhin würden wir uns freuen, wenn wir Sie in Zukunft für weitere Studien kontaktieren dürften. Bitte beachten Sie, dass Ihre Kontaktdaten getrennt von den Befragungsdaten gespeichert werden. Die Anonymität Ihrer Angaben ist gewährleistet.

1. Auswahloption Gewinnspiel
2. Auswahloption Information
3. Auswahloption weitere Studien

Skalierung:

1 „zustimmung“

Kommentar:

Spezieller Fragetyp „Kontaktdaten getrennt erheben“ zur Wahrung der Anonymität der Teilnehmer.

Quellen:

<https://www.soscisurvey.de/help/doku.php/de:create:questions:contact>