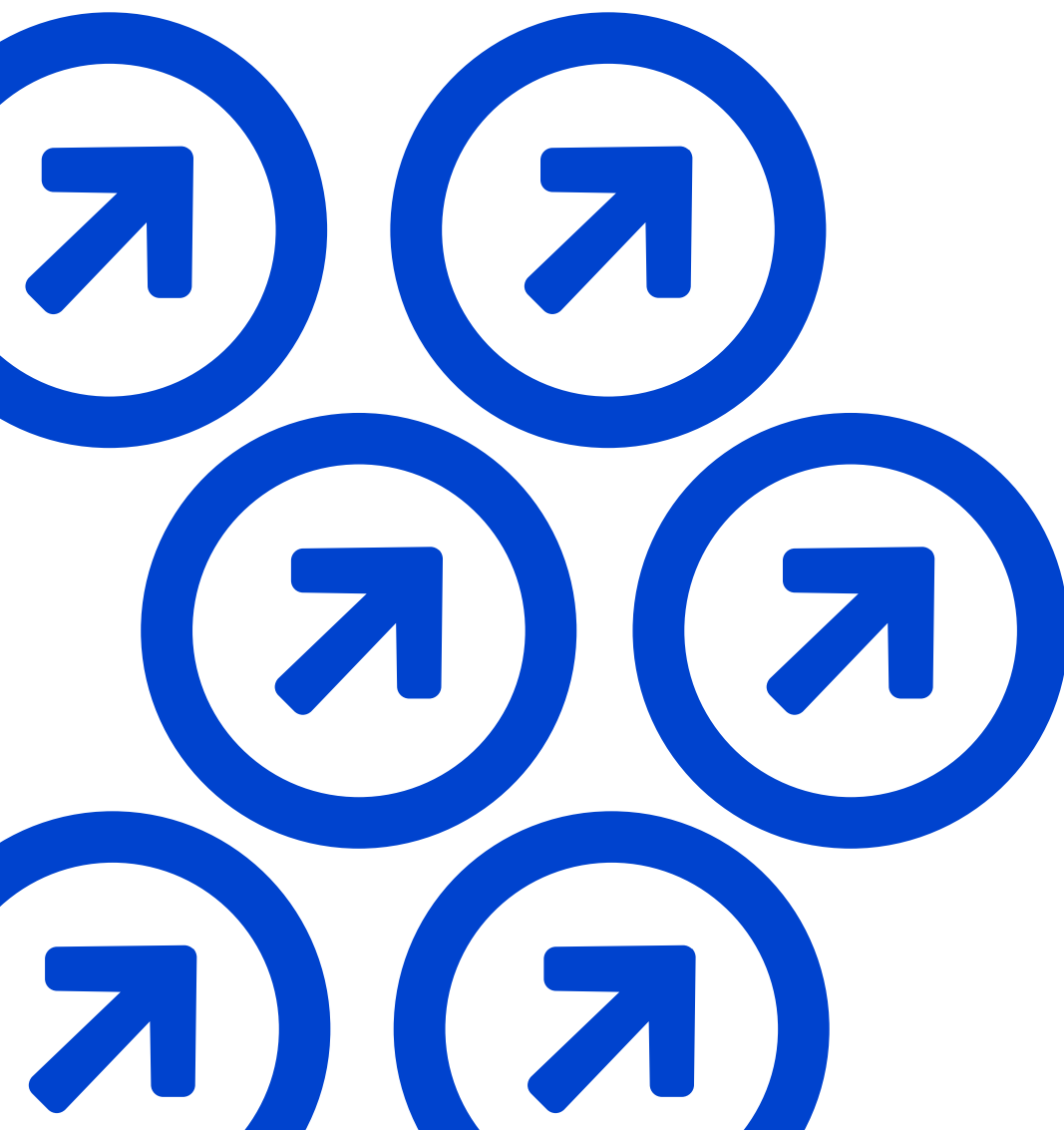


Perspektive der Anwender:innen

Berücksichtigung der Bedürfnisse von Testenden und Nutzenden während der Projektphase und darüber hinaus.

Ein Leitfaden für Projektpartner in KI-Forschungsprojekten.



Gefördert durch:



Bundesministerium
für Wirtschaft
und Klimaschutz

aufgrund eines Beschlusses
des Deutschen Bundestages

Ansprechpartnerin bei Rückfragen

Birgid Eberhardt, GSW Sigmaringen mbH, b.eberhardt@gsw-sigmaringen.de

Annette Hoppe, GSW Sigmaringen mbH, a.hoppe@gsw-sigmaringen.de

Datum

26.08.2025

Version 1.4

Inhalt

1	Perspektive der Anwender:innen in Forschungsprojekten – warum?	5
1.1	Ziel des Leitfadens	5
1.2	Grundlegende Überlegungen	6
1.3	Problemfelder beim Einbezug von Anwender:innen in Forschungsprojekte	7
1.3.1	Anwender:innen	7
1.3.2	Einflussnehmende und Mitbestimmende im Kontext von Anwender:innen	7
2	Projektphasen und Arbeitsschritte mit unmittelbarer Beteiligung von Anwender:innen (User Journey)	10
2.1	Themen während aller Phasen einer Beteiligung	10
2.1.1	Bilder im Kopf – wen adressiert die Lösung?	10
2.1.2	Schriftstücke und Sprache	11
2.1.3	Geschichten erzählen	11
2.1.4	Kommunikationskanäle	12
2.2	Gewinnung Interessierter zur Evaluation geplanter Anwendungen	13
2.3	Evaluation geplanter Anwendungen mit potenziellen Nutzer:innen	13
2.4	Gewinnung von Testenden geplanter Anwendungen	15
2.5	Onboarding von Testenden	16
2.5.1	Pre-Boarding: Vorbereitung der Testbeteiligung	16
2.5.2	Kickoff: Start der Testbeteiligung	17
2.5.3	Onboarding: Begleitung der Testbeteiligung	17
2.6	DropIn, DropOut	19
2.6.1	DropIn	19
2.6.2	DropOut	20
2.7	Exitstrategie, Offboarding	20
3	Projektphasen und Arbeitsschritte indirekter Betroffenheit von Anwender:innen	21
3.1	Erstellung vertrauenswürdiger KI-Anwendungen	21
3.1.1	Einhaltung des AI Acts bzw. der deutschen KI-Gesetze	21
3.1.2	Risikobewertung als Basis für Regulierungen	22
3.1.3	Prüflabel für vertrauenswürdige KI	22
3.1.4	Diskriminierungsfreie Daten als Voraussetzung vertrauenswürdiger KI	23
3.1.5	Transparenzpflicht als Voraussetzung vertrauenswürdiger KI	24
3.2	Erstellung diskriminierungsfreier KI-Anwendungen	24
3.2.1	UX/UI	24
3.3	Materialien für die Gestaltung sozialverantwortlicher KI	27
3.3.1	Kursreihe zur Gestaltung „Sozialverantwortlicher KI“	27
3.3.2	„Faires KI-Prompting - Ein Leitfaden für Unternehmen“	27

Abbildungsverzeichnis

Abbildung 1: Einflussnehmende und Mitbestimmende	9
Abbildung 2: Berührungspunkte von Anwender:innen während der Projektlaufzeit – die User Journey	10
Abbildung 3 Quelle: Internet, anonym	12
Abbildung 4: Risikostufen von KI-Anwendungen (Quelle: u.a.KI-Gesetz)	22
Abbildung 5: Kriterien der VDE SPEC 90012 V 1.0 (en); Übersetzung GSW in Ab- stimmung mit VDE; Quelle: Vertrauenswürdige und ethisch vertretbare KI als Grund- lage für deren Akzeptanz im Rahmen des Forschungsprojektes ForeSight	23
Abbildung 6: Diskriminierungsreihe KI-Anwendungen	24
Abbildung 7: Diskriminierungsbereiche am Beispiel des KI-gestützten Gebäudepfort- ners in ForeSight	26

1 Perspektive der Anwender:innen in Forschungsprojekten – warum?

Der Leitfaden richtet sich an alle Partner von Forschungsprojekten. Er soll als Orientierungshilfe für alle jene Teilprojekte dienen, die mit den Bedürfnissen der Anwender:innen in Kontakt kommen. Diese Bedürfnisse können unmittelbar durch Tests und Evaluation (siehe Kapitel 2, „Projektphasen und Arbeitsschritte mit unmittelbarer Beteiligung von Anwender:innen (User Journey)“) oder mittelbar bei der Entwicklung z. B. von KI-basierten Anwendungen (siehe Kapitel 3, „Projektphasen und Arbeitsschritte indirekter Betroffenheit von Anwender:innen“) relevant sein.

Nicht alle Themen, die im folgenden adressiert werden, sind für alle Projekte oder Teilprojekte gleichermaßen von Interesse.

1.1 Ziel des Leitfadens

Das Ziel dieses Leitfadens ist es, die unterschiedlichen Partner eines Forschungsprojektes dafür zu sensibilisieren, Anwender:innen, die sich zur Mitarbeit und zum Testen in Forschungsprojekten zur Verfügung stellen, nicht nur als Teilnehmer:innen und Datengebende, sondern Mitforschende, die Wissen generieren, zu respektieren. Hierzu gehört, deren Perspektiven und Bedürfnisse bereits im **Vorfeld** von Projektplanungen in diese zu integrieren (Arbeitspakete, finanzielle und personelle Ressourcen) und während der gesamten **Projektlaufzeit** bis hin zu dessen **Ende** im Auge zu behalten und zu berücksichtigen.

Das Ziel eines Forschungsprojektes, dass die Perspektiven von Anwender:innen berücksichtigt, sollte sein, bei Projektende nicht nur evaluierte Ergebnisse vorzuweisen, sondern auch zufriedene Menschen zurückzulassen, die sich ggfs. auf die Beteiligung in weiteren Forschungsprojekten freuen!

Gleichzeitig sollen durch den Leitfaden Unternehmen dafür sensibilisiert werden, die Perspektive und Anforderungen von Anwender:innen in allen Phasen des Produktlebenszyklus von digitalen und durch KI intelligent gestalteten Lösungen zu berücksichtigen. Denn Beispiele zeigen, dass Themen, die während Projektphasen von Forschungsprojekten relevant sind, häufig auch für den „normalen“ Routineeinsatz von Produkten eine Rolle für die Akzeptanz und Nutzungsbereitschaft spielen.

Die Autorinnen dieses Leitfadens haben 20 Jahre Erfahrungen in der Begleitung, Durchführung und in der Beteiligung von Forschungsprojekten. Konkret in diesen Leitfaden eingeflossen sind Erkenntnisse aus dem unmittelbar zurückliegenden Forschungsprojekt ForeSight¹ und den aktuellen Projekten SECAI² und SmartLivingNEXT/ForeSightNEXT³.

¹ <https://foresight-plattform.de/>

² <https://www.secai-energy.de/>

³ <https://smartlivingnext.de/smartlivingnext-foresightnext/>

1.2 Grundlegende Überlegungen

Zu Ausschreibungen für anwendungsnahe Forschungsfelder gehört die Notwendigkeit, in einem Forschungskonsortium neben Forschungs- und Umsetzungspartnern **Anwendungspartner** für angedachte Lösungen einzubeziehen. Forschungspartner aus der **Wissenschaft** oder **Wirtschaft** profitieren davon, ihre während der Projektlaufzeit erforschten und entwickelten Lösungen durch Anwender:innen aus angedachten Zielgruppen bewerten und testen zu lassen.

Anwendungspartner können gewerbliche Unternehmen (z. B. Wohnungsunternehmen oder Dienstleister aus der Pflege) oder auch solche aus einem nicht-gewerblichen Bereich (z. B. ehrenamtliche Nachbarschaftshilfe) sein.

Die Aufgabe von Anwendungspartnern im Konsortium besteht darin sicherzustellen, dass

- ↪ Lösungen passgenau für existierende Problemstellungen entwickelt werden.
- ↪ Rechtliche, ethische und soziale Rahmenbedingungen eingehalten werden.
- ↪ Ggfs. Daten zur Entwicklung von Algorithmen bereitgestellt werden.
- ↪ Testungen und Evaluationen in realen Arbeits- und Lebensumständen durch Zugang zu Anwender:innen durchgeführt werden.
- ↪ Realistische Geschäftsmodelle entwickelt werden.
- ↪ Nach Projektende ein Transfer in die Praxis gelingen kann.

Als Unternehmen oder Organisation haben Anwendungspartner Interesse an den Forschungsergebnissen und stellen daher den **Zugang zu ihren Kund:innen oder Klient:innen** her. Der geplante und erhoffte Nutzen der Forschung kann dabei primär den beteiligten Unternehmen, ihren Kund:innen oder aber weiteren Nutzergruppen aus dem Umfeld der Kund:innen zugute kommen. Dabei werden für die Evaluation von Lösungsideen und das Testen von prototypischen Anwendungen Menschen eingebunden, die in der Regel forschungsfern sind.

Bsp. Ein Sozialwerk betreibt ambulant betreute Wohngruppen. Anwender:innen im Rahmen eines Forschungsprojektes können neben deren Bewohner:innen dort arbeitende Fachkräfte oder An- und Zugehörige sein.

1.3 Problemfelder beim Einbezug von Anwender:innen in Forschungsprojekte

1.3.1 Anwender:innen

Im Gesetz zu unmittelbar an Forschungsprojekten beteiligten Projektpartnern ist die **Motivationslage** von beteiligten Anwender:innen beim Schreiben eines Forschungsantrages häufig unklar. Anwendungspartner in Forschungsprojekten erklären sich zwar bereit, einen Zugang ggfs. zu deren Daten, zu deren Wohn- und/oder Arbeitsumgebungen, zu deren personeller Ressourcen zu schaffen und sie bei der Erfüllung ihnen zugedachter Aufgaben zu begleiten. Es gibt aber in der Regel keinen Pool von Menschen, die bereits im Vorfeld einer Antragsstellung ihr Einverständnis zu einer Teilnahme erklärt haben.

In Forschungsanträgen finden sich daher Arbeitspakete, die die **Gewinnung von Anwender:innen** für unterschiedliche Aufgaben adressieren, die notwendige Einbauten oder das Vorhalten von digitalen oder sonstigen Infrastrukturen berücksichtigen oder das Bewerten der Lösungsansätze bzgl. ihrer Akzeptanz, Effizienz und Effektivität beinhalten. Beteiligte Menschen haben jedoch auch einen Anspruch darauf, dass es eine **geordnete Rückabwicklung** eines Testsettings gibt, sie durch die Beteiligung keine späteren Nachteile erwarten und ggfs. auch über Ergebnisse informiert werden.

Eine **Anforderungsanalyse** nach Beginn der Umsetzung eines Forschungsprojektes, die u.a. festhält, was sich durch das Einbeziehen von Anwender:innen ergeben könnte, kommt u. U. zu Erkenntnissen, die zu diesem Zeitpunkt schon nicht mehr in die **Modifikation geplanter oder die Kreation neuer Arbeitspakete** einfließen können: personelle, finanzielle und sonstige Ressourcen können nicht mehr in dem erforderlichen Umfang bemessen, eingeworben und verteilt werden.

Daher möchte dieser Leitfaden auch dafür sensibilisieren, die Perspektive der Anwender:innen früh und umfänglich bei der Planung, Durchführung und Abwicklung von Forschungsprojekten zu berücksichtigen.

1.3.2 Einflussnehmende und Mitbestimmende im Kontext von Anwender:innen

Zu berücksichtigen ist, dass in Forschungsprojekte involvierte Anwender:innen nicht automatisch eigenständig für sich selber entscheiden. Ca. 1,3 Millionen Menschen stehen unter gesetzlicher Betreuung⁴. Innerhalb von Familien vereinbaren Unterstützungsbedürftige und Sorgende häufig eine Vorsorgevollmacht, um eine gesetzlichen Betreuung zu umgehen.

- Bei Anwendungen, die Hochaltrige, Bewohner:innen im Betreuten Wohnen, Service Wohnen, mit ambulanter Betreuung oder in stationären Settings

⁴ <https://www.betreunungsrecht.de/betreuung/entwicklung-des-betreunungsrechts-betreunungszahlen/>

betreffen, werden Entscheidungen häufig durch **Angehörige** und/oder **Pflegekräfte** beeinflusst oder (mit-)getroffen.

Von kognitiven Einschränkungen Betroffene oder aus anderen Gründen nicht mehr geschäftsfähige Ältere ⁵ haben ggfs. eine Person mit **Vorsorgevollmachten** (Familienangehörige:r) oder eine **Betreuung** (vom Amtsgericht bestellte Unterstützung) an ihrer Seite.

Sind ambulante oder stationäre Pflegekräfte oder andere angestellte Fachkräfte rund um die Anwender:innen involviert, muss ggfs. die Betriebsleitung bzw. ein Betriebsrat involviert werden.

- Auch jüngeren Anwender:innen steht gegebenenfalls eine gesetzliche bestellte **Betreuung** zur Seite, die ihnen hilft, ihre Alltagsgeschäfte zu koordinieren. Zu einer Betreuung führen z. B. körperliche Behinderung einschl. Schlaganfall, seelische Behinderung/psychische Erkrankungen, Altersdemenz, geistige Behinderung.⁶

Bezogen auf Mietende gibt u. U. einen Betreuer, „der für den Aufgabenkreis Aufenthaltsbestimmung und/oder Wohnungsangelegenheiten bestellt ist, (er) kümmert sich um den Abschluss und die Erhaltung von Mietverträgen, falls nötig, um den Umzug in ein geeignetes Heim, regelt die Kündigung des Mietverhältnisses und die Auflösung des Haushaltes.“ ⁷

Da diese Betreuung häufig eine Vertrauensperson ist, wenden sich Betroffene bei Problemen einer forschungsbedingten Testinstallation zunächst an dieselbe.

- Immer, wenn Minderjährige involviert sind, müssen deren Erziehungsberechtigte involviert werden.

⁵ <https://www.bpb.de/kurz-knapp/lexika/lexikon-der-wirtschaft/19499/geschaeftsfaehigkeit/> (3.2.2025)

⁶ <https://www.kvjs.de/soziales/service-betreunungsrecht/betreunungsstatistik>

⁷ <https://www.berliner-mieterverein.de/magazin/online/mm0422/10-fragen-zur-betreuung-ein-starker-beistand-fuer-den-hilfsbeduerftigen-mieter-022224.htm#9-Kann-auch-der-Vermieter-eine-Betreuung-anregen> (3.2.2025)

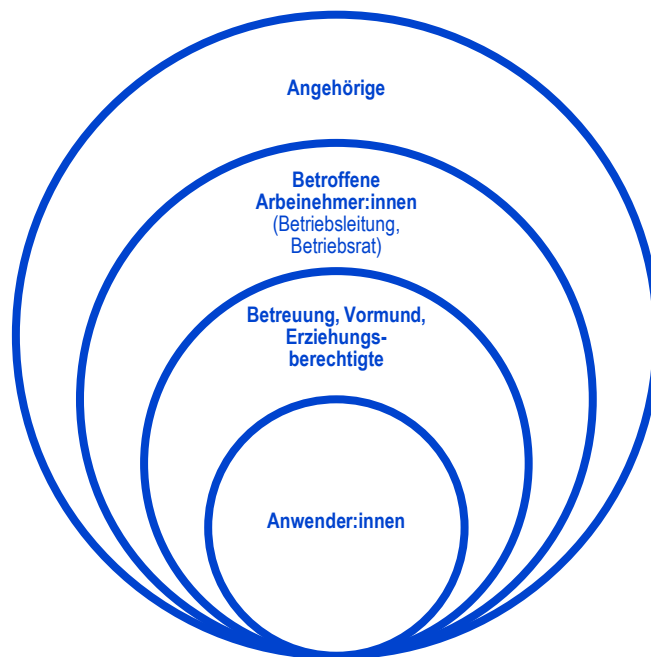


Abbildung 1: Einflussnehmende und Mitbestimmende

⚡ Es gibt Personen unter Betreuung oder Vormundschaft. Je nach Anwendungsfall müssen die hierfür verantwortlichen Personen mitgedacht oder sogar aktiv in ein Testgeschehen, in Befragungen etc. einbezogen werden.

Im weiteren Verlauf der Leitfadens werden nur die Anwender:innen genannt, die Einflussnehmenden und Mitbestimmenden jedoch mitgedacht.

2 Projektphasen und Arbeitsschritte mit unmittelbarer Beteiligung von Anwender:innen (User Journey)

Während eines Forschungsprojektes können Anwender:innen in verschiedenen Stadien und zu unterschiedlichen Aufgabenstellungen unmittelbar Nutzen stiften (Co-Creation). Während manche Menschen nur während bestimmter Projektphasen eingebunden werden, begleiten andere ggfs. ein Projekt durch unterschiedliche Phasen über längere Zeiträume hinweg (User Journey).

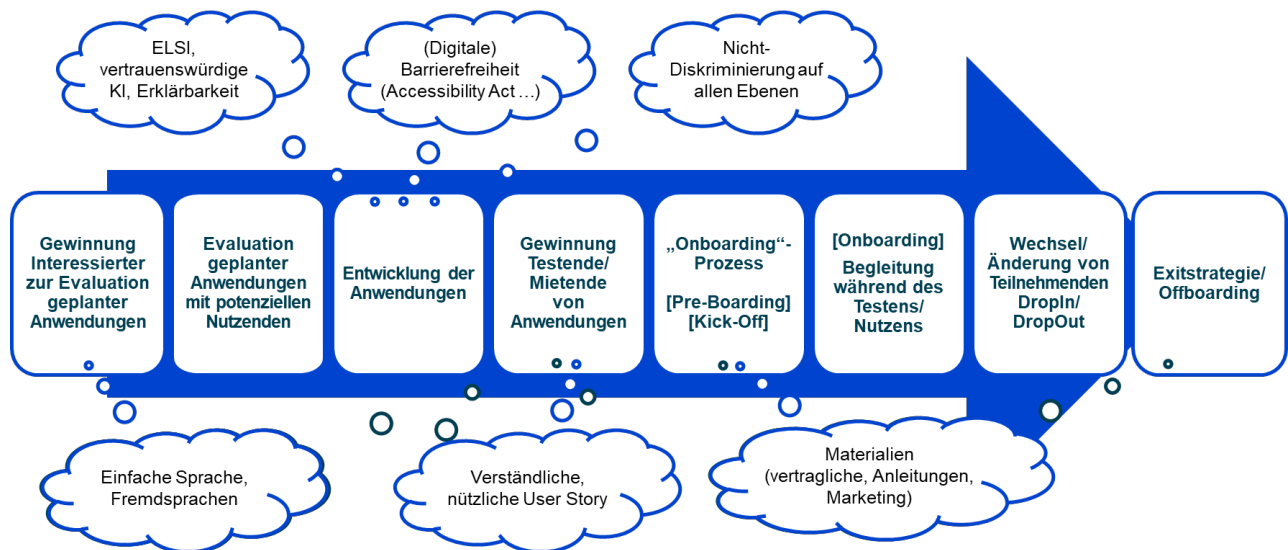


Abbildung 2: Berührungspunkte von Anwender:innen während der Projektlaufzeit – die User Journey

2.1 Themen während aller Phasen einer Beteiligung

2.1.1 Bilder im Kopf – wen adressiert die Lösung?

Ein bekanntes Phänomen in der Produktentwicklung und Forschung ist, sich auf Teil-Zielgruppen zu konzentrieren, ohne diese explizit zu benennen.

Bsp.

- Lösungen für Einfamilienhausbesitzer taugen nicht zwingend für Mehrfamilienhäuser der Wohnungswirtschaft.
- Apps für effizientes Heizen motivieren selten Haushalte, die Miete und Heizung als Transferhilfe bekommen.
- Ältere, digitalferne Personen können disruptiven Prozesse, die auf Apps beruhen, nicht nutzen und werden ausgeschlossen/abgehängt.
- Deutschland ist ein Multi-Kulti-Land mit Menschen, die nicht zwangsläufig der deutschen Sprache mächtig sind und sich stark von mittelalten Menschen/Männern kaukasischen Typs unterscheiden.

- Der Anteil alter und hochaltriger Menschen mit ihrer wachsenden Diversität über ein großes Altersspektrum hinweg nimmt zu (siehe hierzu den Achten Altersbericht der Bundesregierung⁸).

Das Abbild dieser „inneren“ Zielpersonen/-gruppen findet sich während der Phase von Forschungsprojekten in klichehaften Personas (z. B. Oma Erna), User Stories und Bildern wieder und nimmt ihren Lauf in der späteren Produktgestaltung, in Marketingunterlagen etc.

- 👍 Klare Zielgruppen zu definieren, die ggfs. deutlich definierte Personen begründet ausschließen, ist legitim.
- 🗨️ Geschieht dies jedoch nicht bewußt und explizit, kommt es leicht zu ungewollten Diskriminierungen.
- 🔍 Zielgruppen genau beschreiben, inklusive explizit nicht berücksichtigter Personengruppen.

2.1.2 Schriftstücke und Sprache

Ca. ein Viertel der in Deutschland lebenden Bevölkerung hat einen Migrationshintergrund. Die sprachliche Fähigkeiten dieser Personengruppe sind sehr divers⁹.

Ca. 6,2 Millionen Menschen oder 12,1 Prozent der erwerbsfähigen Bevölkerung können in Deutschland nicht oder nur unzureichend lesen und schreiben. Bei weiteren 10,6 Millionen Menschen oder 20,5 Prozent der Erwachsenen tritt fehlerhaftes Schreiben selbst bei gebräuchlichen Wörtern auf¹⁰. Lesen und Schreiben als Grundfertigkeiten können nicht als gegeben vorausgesetzt werden.

In Forschungskontexten geht es häufig um komplexe Sachverhalte, die selbst in der interprofessionellen und interdisziplinären Kommunikation unter den Projektpartnern nicht zwingend Beachtung finden.

- 🔍 Gegenüber beteiligten Anwender:innen ist es umso wichtiger, Fachsprachen, Anglizismen und Abkürzungen zu vermeiden.

2.1.3 Geschichten erzählen

„Der Wurm muss dem Fisch schmecken und nicht dem Angler.“¹¹ Dieses vielleicht nicht mehr ganz dem Zeitgeist entsprechende Zitat beschreibt, wie wichtig es ist, die Wünsche der potenziellen Anwendenden/Kunden in User Stories in den Vordergrund zu stellen und nicht die eigenen Ideen.

⁸ Vollständiger Bericht: https://www.neunter-altersbericht.de/fileadmin/Redaktion/Bericht_Broschuere_Deckblaetter/neunter-altersbericht-bundestagsdrucksache_final.pdf, Kurzfassung: https://www.neunter-altersbericht.de/fileadmin/Redaktion/Bericht_Broschuere_Deckblaetter/Broschuere_Neunter_Altersbericht_barrierefrei.pdf

⁹ https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/_inhalt.html

¹⁰ <https://www.mein-schlüssel-zur-welt.de/de/helfen/zahlen-und-fakten/zahlen-und-fakten.html>

¹¹ Frei nach Prof. Dr. Helmuth Thoma

Um dies zu erreichen, beinhaltet eine gute Geschichte immer ein Was, ein Wie und ein Wozu.

Fünf Faktoren tragen maßgeblich dazu bei, dass eine Geschichte oder auch User-Story erfolgreich ist.

Diese sind:

- Relevanz: Ist es wichtig für den Anwendenden?
- Authentizität: Lebensnahe Protagonisten
- Emotionalität: Löst es ein Problem?
- Visualisierung: Bildhafte Sprache
- Klarheit: Einfache Sprache, kein „Fachchinesisch“

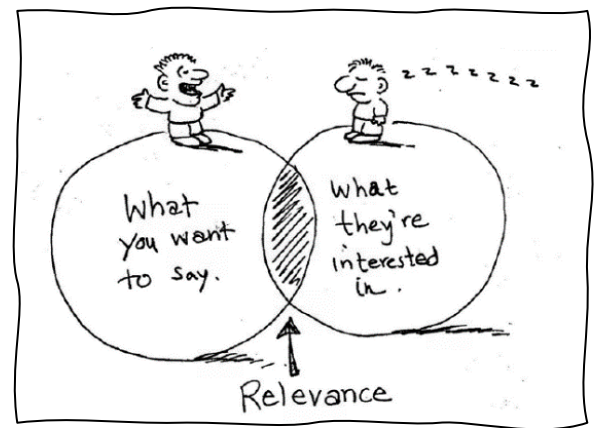


Abbildung 3 Quelle: Internet, anonym

- 🗨️ Use Cases beschreiben Anwendungen aus Sicht der Implementierung, der Technik. Sie sind in der Kommunikation mit Anwender:innen nicht hilfreich.
- 👍 Gegenüber Anwender:innen helfen gut ausformulierte, nachvollziehbare Geschichten, User Stories, die einen konkreten, erkennbaren Nutzen beschreiben.
- 🔧 Hilfreich sind einfach nachvollziehbare Geschichten mit klarem Bezug zu den Anwender:innen und einem erkennbarem Nutzen.

2.1.4 Kommunikationskanäle

Anwender:innen benötigen verlässliche Ansprechpartner und Kommunikationswege.

- 🔧 Die Kommunikation mit den Testenden und die Begleitung während der Beteiligung im Forschungsprojekt sollte über vertraute, nicht wechselnde Personen („one face to the customer“) erfolgen.
- 🔧 Bei der Wahl der Kommunikationswege sind die Gewohnheiten und Fähigkeiten der Beteiligten zu berücksichtigen. Auch in KI- oder Digitalprojekten ist ein Aushang, ein Anschreiben oder eine direkte Ansprache manchmal hilfreicher als ein digitaler Kommunikationsweg (Mieter-App, Email etc.).
- 🔧 Ziele
 - Sozialwissenschaftlich „saubere“ Kommunikation
 - Keine BIAS
 - Kein Vertrauensverlust gegenüber dem vertrauten Anwendungspartner!

2.2 Gewinnung Interessierter zur Evaluation geplanter Anwendungen

Während der Formulierung von Forschungsanträgen sind Anwendungsfälle und die dafür notwendigerweise zu erfüllenden Anforderungen in der Regel noch nicht konkret beschrieben. Und auch wenn Anwendungspartner bei der Konkretisierung involviert werden, kann es sinnvoll sein, spätere mögliche Anwender:innen vor dem Beginn der Umsetzung dieser Anwendungsideen nach deren Akzeptanz zu befragen.

Für die Evaluation von Ideen können Menschen angesprochen werden, die später auch in Tests eingebunden werden. Es ist aber auch möglich mit Fokusgruppen zu arbeiten, die definierte Anforderungen bzgl. ihrer sozio-demografischen Ausprägung (Alter, Geschlecht, Bildungsgrad, ...) erfüllen sollen, um spätere Nutzer:innen zu repräsentieren.

-  Use Cases und technische Beschreibungen sind keine Basis, um Anwender:innen einzuladen. Fachsprache und -ausdrücke sind nicht hilfreich.
-  Geschichten und Fragestellungen, die mögliche Nutzer:innen in ihrer Lebenswelt abholen, können ihre Neugier wecken und sie zum Mitmachen motivieren.
-  Schriftliche, informative Einladung an potenzielle Anwender:innen oder Verteiler von Fokusgruppen; Aufsuchen von Gruppen (z. B. Senioren an -nachmittagen, in -zentren); Einladung über Vertrauenspersonen (z. B. Ehrenamtliche)

Potenzielle Anwender:innen, die sich bereit erklären, sich User Stories anzuhören und sich in die Geschichten und Lösungsansätze hineinzudenken, sind in der Regel von sich aus motiviert.

2.3 Evaluation geplanter Anwendungen mit potenziellen Nutzer:innen

Die Evaluation von Anwendungsideen und die Feststellung ihrer Akzeptanz können hilfreich sein, um diese zu schärfen, Prioritäten für die Realisierung zu setzen oder zu verschieben und ggfs. bereits ausformulierte Lösungsansätze zu verwerfen.

Um Lösungsansätze zu diskutieren und ein begründetes Votum erarbeiten zu können, sind gut ausformulierte User Stories unumgänglich. Neben erzählenden Bestandteilen können Sketche, Bilder oder Videos hilfreich sein, besser in Problemstellungen, Lösungen und Nutzungserfahrungen einzutauchen. Diese sollten die Lebenswirklichkeit der beurteilenden Personen aufgreifen.

Anwendungsfälle erforschende und/oder umsetzende Konsortialpartner sollten bereit sein, die Ergebnissen einer solchen Evaluation ernst zu nehmen.

-  Use Cases und technische Beschreibungen oder „Vorteile“ sind keine Basis, um Anwender:innen mitzunehmen und in Lösungseinsätze einzuführen.

- 👍 Geschichten und Fragestellungen, die möglichen Nutzer:innen den Nutzen von geplanten Lösungsansätzen verdeutlichen, helfen ihnen, sich in diese hinein zu denken, sie zu bewerten, weiterzudenken und sachliche Kritik anzubringen.
- 🔗 Powerpoint oder gedrucktes Material mit Scribbles, Videoclips, welche die User Story verdeutlichen und untermalen; Bewertung über Noten; Bewertung über Emojis; Bewertung durch Fragebogen; Leitfaden gestützte Interviews; offene Diskussion.

Bsp. In ForeSight ergaben sich durch die Evaluation von User Stories folgende Erkenntnisse und Veränderungen in der Umsetzung:

1. Eine KI-basierte Gesichtserkennung wurde im privaten Umfeld wertgeschätzt; das Betreten der Wohnung durch unbekannte Personen wie Lieferdienste oder Handwerker, wurde abgelehnt.
 - ▶ Die Entwicklung konzentrierte sich auf die Umsetzung des Anwendungsfalls, bekannten Personen mittels einer KI-basierten Gesichtserkennung den Zutritt zu ermöglichen.
2. Ein KI-basierter Ernährungsassistent, der Vorlieben und Beeinträchtigungen unterschiedlicher Familienmitglieder berücksichtigen sollte, fand keine Zustimmung. Die Mitglieder der Fokusgruppe entwickelten eine eigene Idee, nämlich die eines Assistenten zum Vermeiden von verdorbenen Lebensmitteln.
 - ▶ Der Anwendungsfall wurde auf Basiserkenntnisse reduziert und diese nur im Labor getestet.

2.4 Gewinnung von Testenden geplanter Anwendungen

Um während der Entwicklungsphase Anwender:innen für ein kooperatives Design und nach der Fertigstellung von nutzbaren Artefakten Testpersonen im Kontext ihres Lebens- und/oder Arbeitsraums zu gewinnen, bedarf es umfänglicherer Informationsmaterialien. Diese sollten, analog zu 2.1, gut erzählte User Stories mit einem nachvollziehbaren Nutzen enthalten.

Eine wesentliche Unterscheidung in der Motivation der Beteiligung liegt darin begründet, wie hoch der Grad der Freiwilligkeit und der intrinsischen Motivation ist.

Bsp. Sollen spätere Endverbraucher:innen für den Test von digitalen Heizkörperthermostaten in Eigenheimen gewonnen werden, kann bei deren Zustimmung eine intrinsische Motivation und ein eigenes Interesse vorausgesetzt werden.

Sollen hingegen Mietparteien digitale Heizkörperthermostaten testen, die über ihren Vermieter als Anwendungspartner eines Forschungsprojektes in ihre Wohnungen gelangen, kann von einer intrinsischen Motivation und einem eigenen Interesse nicht automatisch ausgegangen werden.

Bsp. Wollen Endverbraucher:innen einen sicheren Rauchwarnmelder nutzen und wählen sie ein Produkt nach entsprechender Information nach ihren Bedürfnissen, liegt bei deren Zustimmung eine intrinsische Motivation und ein eigenes Interesse zugrunde.

Verbaut hingegen ein Vermieter neuartige, sichere Rauchwarnmelder, kann von einer intrinsischen Motivation und einem eigenen Interesse der Mietparteien nicht automatisch ausgegangen werden.

Je nach Motivationslage muss die Ansprache anders gestaltet werden. Die Motivationslage sollte auch in der Bewertung der Testergebnisse einfließen, um diese nicht zu verzerren.

Es kann hilfreich sein, die Motivation für die Beteiligung an Tests oder einer späteren Evaluation durch Aufmerksamkeiten zu steigern:

- Gutes Catering bei Veranstaltungen
- Aussicht auf Weiternutzung oder Eigentumsübertragung nach Testende
- In die Testphase integrierte Wettbewerbe mit erstrebenswerten Gewinnen
- Finanzielle Anreize (in geförderderten Projekten nicht förderfähig ...)

 Use Cases und technische Funktionsbeschreibungen oder „Vorteile“ sind keine Basis, um Anwender:innen für eine Testteilnahme zu gewinnen.

 Geschichten und Fragestellungen, die möglichen Testenden den Nutzen von geplanten Lösungsansätzen verdeutlichen, helfen ihnen, diese im Rahmen ihrer Lebenswirklichkeit zu bewerten, weiterzudenken und sachliche Kritik anzubringen.

Die potenziellen Anwender:innen sollten über das Testsetting informiert werden,

über Erwartungen bzgl. ihrer Bewertung, über Möglichkeiten der Unterstützung, über ggfs. vorgesehene Befragungen, über die Abwicklung eines vorgesehenen Endes der Testperiode.

2.5 Onboarding von Testenden

Wenn Anwender:innen gefunden wurden, die sich für eine Teilnahme an den Tests im Rahmen eines Forschungsprojektes oder auch einer Produktevaluation bereitfinden, sind Prozessschritte zu beschreiben, mit Ressourcen zu versehen und zu organisieren, die sich am Onboarding-Prozess aus dem Personalmanagement orientieren können.

2.5.1 Pre-Boarding: Vorbereitung der Testbeteiligung

Eine allgemeine Aufklärung über das Testsetting, ggfs. notwendige Installationen und Rahmenbedingungen, die Testdurchführung und damit ggfs. notwendige Anpassungen an gewohnte Abläufe, die Bewertung der Nutzung der Exponate und eine Beschreibung des Testendes sollten in einer Einverständniserklärung zusammengefasst werden, die von den Teilnehmenden unterzeichnet wird.

Werden durch das Testen persönliche Daten der Anwender:innen oder aus ihrem Umfeld erhoben und verarbeitet, muß im Rahmen eines Informed Consent genau erklärt werden, welche Daten zu welchem Zweck erhoben werden, wo sie gespeichert und zu was sie verarbeitet werden. Worüber konkret im Informed Consent informiert werden muss, ist von Projekt zu Projekt und Anwendungsfall zu Anwendungsfall unterschiedlich. Es ist sinnvoll, hierzu Datenschutzbeauftragte hinzu zu ziehen. Jeder Informed Consent muss, um gültig zu sein, verständlich formuliert sein, ggfs. in der Muttersprache der Beteiligten. Seine Beschreibung ist bindend: Im Laufe der Tests dürfen weder die Datenaufnahme erweitert noch ihre Weiterverarbeitung ausgedehnt werden. Es sollte auch beschrieben sein, was am Ende des Projektes mit den Daten geschieht. Testende Anwender:innen unterzeichnen diesen mit ihrer Unterschrift (siehe auch Kapitel „1.3.2 Einflussnehmende und Mitbestimmende im Kontext von Anwender:innen“).

Den Testenden werden Ansprechpartner für die Testphase benannt, Möglichkeiten des Supports sind bekannt.

-  Das Versprechen der Einhaltung einer Datenschutzgrundverordnung/DSGVO allein weckt kein Vertrauen!
-  Generelle Einverständniserklärung mit Testablauf (siehe auch Kapitel „1.3.2 Einflussnehmende und Mitbestimmende im Kontext von Anwender:innen“ und Kapitel „2.1 Themen während aller Phasen einer Beteiligung“)
-  Informed Consent (siehe auch Kapitel „1.3.2 Einflussnehmende und Mitbestimmende im Kontext von Anwender:innen“)

- ↪ Alle Dokumente sollten in verständlicher Sprache, ggfs. zusätzlich in Muttersprachen der Beteiligten, bereit stehen: Beteiligte sollten wirklich wissen und verstehen, auf was sie sich unter welchen Bedingungen einlassen.
- ↪ Der Support ist geklärt; Umfang je nach Anwendungsfall klären (je größer der Eingriff in Alltagsgewohnheiten, desto umfänglicher müssen Supportzeiten sein).
- ↪ Ohne formelle Zustimmung keine Aufnahme in Tests

2.5.2 Kickoff: Start der Testbeteiligung

Je nach Anwendungsfall ergeben sich mit dem ersten Nutzungstag für die Anwender:innen neue oder veränderte Abläufe, die begleitet, ggfs. noch einmal erklärt und durch Anleitungen unterstützt werden müssen.

- ↪ Ab dem ersten Tag der Testphase müssen Anleitungen zur Nutzung bereitstehen. Oft ist ein QuickGuide bzw. QuickStart hilfreich.
- ↪ Für Rückfragen muss die Erreichbarkeit von persönlicher Unterstützung (Support) gewährleistet und bekannt sein. Wichtig sind hierbei die Erreichbarkeit auch außerhalb „normaler“ Geschäftszeiten (Wochenende, Feiertage, ggfs. sogar Abende).
- ↪ Bei der Unterstützung in dieser Phase ist es außerdem hilfreich,
 - Bereits bekannte Kontaktpersonen
 - Ein ggfs. vorhandenes Unterstützungsnetzwerk einzubinden.
- ↪ Auftretende Schwierigkeiten und Verbesserungsvorschläge sollten dokumentiert und allen im Support und in der Entwicklung Beteiligten bekannt gegeben, Lösungen gegenüber Betroffenen wertschätzend kommuniziert werden (keine heimlichen Funktionserweiterungen, kein stilles Bugfixing im Hintergrund).

Oben genannte Punkte gelten umso mehr, je unmittelbarer die zu testenden Anwendungen in den Alltag/die Lebenswelt der Testenden eingreifen (Wärme in den Wohnungen, warmes Wasser, Alarmer durch intelligente Assistenz, ...).

2.5.3 Onboarding: Begleitung der Testbeteiligung

Der Onboarding-Prozess beschreibt die Begleitung während der gesamten Testphase. Hierzu gehören Reaktionen auf Fragen oder auftretende Problemen; oft ergibt sich im Nutzungsprozess der Bedarf von Optimierungen oder Anpassungen an persönliche Lebensumstände. Wichtig sind diese Begleitungen vor allem dann, wenn das Testsetting „geschuldete“ Leistungen betrifft.

Bsp. Wärme in Mietwohnungen ist eine geschuldete Leistung der Vermietung. Kommt es auf Grund von Testsettings in der Wohnung oder im Keller zu Ausfällen der Wärmeerzeugung, -verteilung oder -übertragung und die Räume

lassen sich nicht auf die geschuldete Mindestwärme erwärmen, haben Mietparteien das Recht auf Mietminderung.

Je nach Dauer einer Testphase treten Gewöhnungseffekte ein – oder eben auch auf Grund der Kürze der Testphase nicht. Und je nach Anwendungsfall müssen manche Anwendungen periodisch neu „erlernt“ werden.

Bsp. Mietparteien neigen dazu, vom Ende einer Heizperiode bis zum Beginn der nächsten erlernten Verhalten zu vergessen.

Die Schritte für die Begleitung von Anwender:innen in einem Forschungsprojekt empfehlen sich auch bei der Begleitung von Anwender:innen im Realbetrieb, die ohne eigene Motivation mit neuen Lösungsansätzen, z. B. als Mietende, in Berührung kommen.

Bsp. Die Einführung neuer Rauchwarnmelder von Techem durch die Vonovia führte durch ungenügende Kommunikation und Begleitung zu Protesten.

-  Anwender:innen nur im Rahmen der Kickoff-Phase intensiv zu betreuen ist nicht ausreichend.
-  Eine anfängliche Unterstützung reicht vor allem dann nicht aus, wenn Anwender:innen auf Grund der Periodizität der Nutzung eine Anwendung nicht durchgängig nutzen.
-  Begleitete Anpassungen und umgesetzte Optimierungen resultieren in einer bereitwilligeren und intensiveren Nutzung und einer höheren Akzeptanz (und ggfs. Zahlungsbereitschaft).
-  Auch nach dem KickOff muss während der gesamten Testphase die Betreuung und Begleitung der Anwender:innen organisiert und gewährleistet sein und im Fall einer eher periodischen Nutzung den Anwender:innen ggfs. aktiv näher gebracht werden.

Zur Testphase gehören in der Regel die Evaluationen der Effizienz, der Akzeptanz und manchmal der Zahlungsbereitschaft. Die Bewertung statischer Daten kann in der Regel ohne die Beteiligung der Anwender:innen erfolgen. Um die Akzeptanz oder Zahlungsbereitschaft zu bewerten, können unterschiedliche Methoden zum Einsatz kommen.

-  Nicht alle Evaluationsmethoden sind für alle Anwender:innen gleichermaßen geeignet (z. B. online-Fragebögen in deutscher Sprache).
-  Direkte Gespräche im Testsetting oder eine Kombination unterschiedlicher Methoden (Interviews, begleitetes Ausfüllen von Fragebögen, verschiedene Bewertungsverfahren) ergeben ein deutlicheres Bild von dem erlebten Nutzen, Störungen und einer in die Zukunft gerichteten Akzeptanz der Lösung.
-  Befragungen durch unterschiedliche am Forschungsprojekt beteiligte Partner mit wechselnden Personen sind schwierig für eine vertrauensvolle Evaluation. Der Wunsch unterschiedlicher Projektpartner, unmittelbar mit den

Anwender:innen in Kontakt zu treten, stören (siehe Kapitel „2.1.4 Kommunikationskanäle“).

- 👍 „On face to the Costumer“ hat sich als vertrauensstiftend und die Evaluation fördernd erwiesen (siehe Kapitel „2.1.4 Kommunikationskanäle“).
- 🔧 Alle Evaluationen müssen geplant werden. Nicht statische Evaluationen sollten mit einem großzügigen Zeitbudget versehen und von einem sozialwissenschaftlich/soziologisch geschultem Partner für alle Beteiligten gebündelt geplant, organisiert und durchgeführt werden.

2.6 DropIn, DropOut

Aus unterschiedlichen Gründen können während einer Test- oder realen Nutzungsphase Anwender:innen ausscheiden oder neu hinzukommen.

2.6.1 DropIn

Anwender:innen können sich nach Beginn der Testphase entscheiden, sich an Tests zu beteiligen.

Bsp. Durch die Begeisterung von beteiligten Anwender:innen fühlen sich Parteien, die zunächst nicht für eine Beteiligung gewonnen werden, motiviert, sich noch zu beteiligen.

Bsp. Durch Mieterwechsel kommen „neue“ Mietparteien hinzu.

- 🔧 Prozess für die Aufnahme neuer Anwender:innen in einem neuen Kontext, mit einer neuen Ausstattung. Hierzu ist der gesamte Onboarding-Prozess zur Verfügung zu stellen; ggfs. wird zusätzliche Hardware benötigt, ggfs. weitere Lizenzen, zusätzliche Personal- und Finanzressourcen.
- 🔧 Prozess für die Aufnahme neuer Anwender:innen in eine vorhandene Ausstattung, die bereits vorher im Einsatz war, von der ggfs. bereits Daten gesandt wurden. Neben dem Onboarding der Anwender:innen (siehe oben) ist sicherzustellen, dass neue Daten von bereits vorhanden alten getrennt werden oder aber zu klären, unter welchen Bedingungen sie gemeinsam ausgewertet werden können.
- 🔧 Die Frage, ob und welchen Rahmenbedingungen neue Anwender:innen aufgenommen werden können, sollte klar sein.
- 🔧 Nicht nur im Realbetrieb ist sicherzustellen, dass neue Anwender:innen nicht mit von vorherigen Nutzer:innen vorgenommene Einstellungen „konfrontiert“ werden.

2.6.2 DropOut

Beim Ausscheiden muss geklärt sein, ob (noch) personalisierte Daten vorliegen, wie mit ihnen verfahren wird und wie Anwender:innen deren Verarbeitung (z. B. Löschung) bestätigt werden kann.

Zudem muß geklärt sein, ob es im Testsetting zu einer vorgezogenen Existenzstrategie mit einem Rückbau kommt, ob Ausscheidungsgründe evaluiert werden.

2.7 Exitstrategie, Offboarding

Jede Testphase findet spätestens mit Ende der Projektlaufzeit ein natürliches Ende. Die Exitstrategie beschreibt, wie die Offboarding-Phase geplant und für die Anwender:innen als gestalteter Prozess wahrgenommen werden kann.

Leider hat sich in der Vergangenheit in einer Vielzahl von Projekten gezeigt, dass dieser Phase zu wenig planerische Aufmerksamkeit und in der Folge zu wenig Ressourcen zugeordnet wurden.

Nicht immer ist zu Beginn einer Projektplanung klar, wie ein Projekt konkret enden wird.

Bsp. Technische Artefakte inklusive Einbauten können

1. Inklusive des Testsettings weiterbetrieben werden,
2. Unabhängig vom Testsetting weiterbetrieben werden,
3. Rückgebaut werden.

Bleiben getestete Artefakte in der Nutzung, muss vor dem Übergang in einen Realbetrieb sichergestellt sein, ob und wie lange während der Testphase eingebundene Services weiterlaufen und welche Auswirkungen ihr Fehlen hat. Zudem muss sichergestellt werden, welchen Aufgaben und Kosten auf die zukünftigen Betreiber zukommen, wie deren Prozesse für den laufenden Betrieb und Kernprozesse (wie beispielsweise einen Wechsel der Mietpartei) aussehen müssen und wie sie erfüllt werden können.

Im Falle eines notwendigen Rückbaus ist bereits beim Einbau zu berücksichtigen, was für Folgen der Ausbau für die Umgebung hat.

Bsp. Beim Einbau einer Smart Home-Umgebung in einer Mietwohnung stellt sich die Frage, wo und wie Sensoren verbaut werden. Fragen, die bereits beim Einbau geklärt sein müssen, sind, ob bei der Entfernung bleibende Schäden an Wänden, Tapeten oder Türrahmen zurückbleiben, die ggfs. der Mietpartei später zur Last gelegt werden können.

In die Planungen der Exitstrategie sollte auch die Abschiedskommunikation mit Anwender:innen einfließen. Werden sie an Erkenntnissen beteiligt? Bekommen sie ein Dankschreiben?

3 Projektphasen und Arbeitsschritte indirekter Betroffenheit von Anwender:innen

Es gibt Projektphasen, in denen Anwender:innen zwar nicht unmittelbar eingebunden sind, ihre Bedürfnisse und Bedarfe jedoch von beteiligten Konsortialpartnern bei der Erstellung der KI-Anwendungen mitgedacht werden müssen.

3.1 Erstellung vertrauenswürdiger KI-Anwendungen

Vertrauenswürdige KI ist die Basis für diskriminierungsfreie KI-Anwendungen. Deren Umsetzung und die Berücksichtigung der speziellen KI-Gesetze und -Normen muss im Entwicklungsprozess, ggfs. sogar in Unternehmensstrukturen, angelegt sein.

3.1.1 Einhaltung des AI Acts bzw. der deutschen KI-Gesetze

Die Notwendigkeit der Einhaltung geltender Gesetze und Verordnungen findet sich u.a. bei der Erhebung, Speicherung und Übermittlung von Daten, in deren Verarbeitung, bei der Entwicklung zu KI-Services oder -Anwendungen. Die Umsetzung des geltenden Rechtsrahmen innerhalb der technischen Entwicklungen zum Schutz von Anwender:innen ist notwendig und sollte auf Nachfrage erklärt werden können. Betroffen sind Anwender:innen mehr oder weniger unmittelbar bei:

- Der Entwicklung vertrauenswürdiger KI
- Der Erfüllung der ESG (Ethik, Fairness und Umweltschutz)
- Der Erfüllung allgemeiner gesetzlicher Vorgaben (z. B. Datenschutzgrundverordnung/DGSVO ...)
- Die Erfüllung konkreter gesetzlicher Vorgaben (z. B. Medical Device Regulation/MDR und Medizinprodukte-Durchführungsgesetz/MPDG ...)
- Der Erfüllung von Standards (z. B. ISO 42001¹², VDE SPEC 90012 V1.0)
- Einer ggfs. notwendige Zertifizierung

Der europäische AI ACT trat am 1. August 2024 in Kraft mit dem Ziel, die Grundrechte der Menschen im Einflußbereich der EU zu schützen. Die Umsetzung in nationales Recht hat innerhalb von zwei Jahren zu erfolgen (April 2026).

Die Einhaltung der internationalen Regulierungen aus dem AI Act¹³ bzw. deren Umsetzung in deutsches Recht ist die Grundvoraussetzung zur Erstellung vertrauenswürdiger KI-Anwendungen. Einen Anhaltspunkt für deren Kriterien kann die VDE SPEC 90012 V1.0¹⁴ geben.

¹² Siehe z. B. Beitrag bei der DEKRA zum Rahmen für den sicheren Einsatz von KI-Systemen in der Wertschöpfungskette: <https://www.dekra-certification.de/de/iso-42001-potenziale/>

¹³ <https://aistandardshub.org/>

¹⁴ <https://www.vde.com/resource/blob/2337818/a24b13db01773747e6b7bba4ce20ea60/vcio-based-description-of-systems-for-ai-trustworthiness-characterisationvde-spec-90012-v1-0--en--data.pdf>

3.1.2 Risikobewertung als Basis für Regulierungen

Die Basis für alle Regulierungen ist eine Risikobewertung. Die Einstufung in einer der Risikoklassen beruht auf einer Selbsteinschätzung.

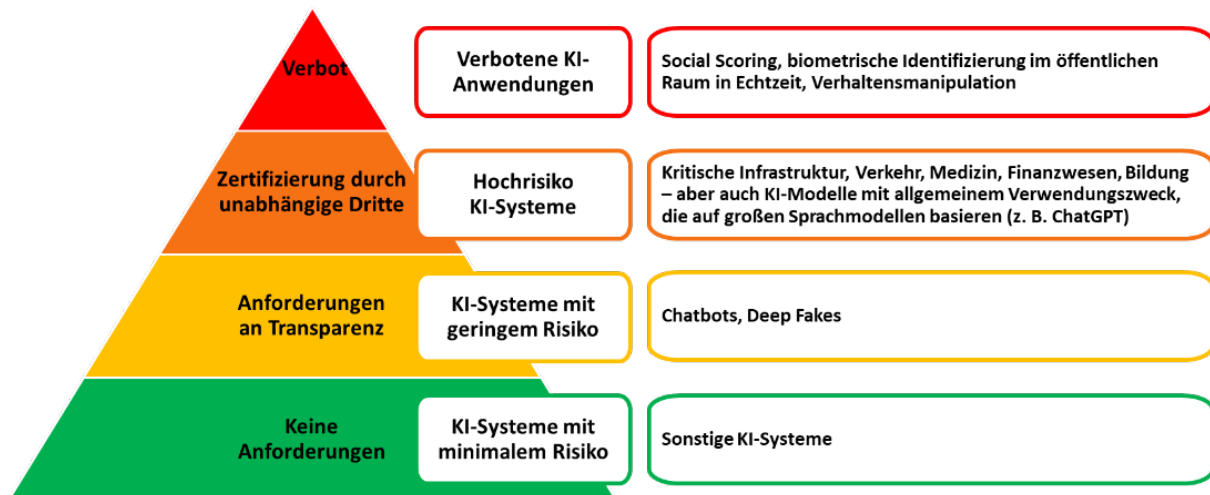


Abbildung 4: Risikostufen von KI-Anwendungen (Quelle: u.a. KI-Gesetz ¹⁵)

Aus der Einstufung ergibt sich zum Beispiel zwingend die Notwendigkeit, Ergebnisse aus KI-Services erklären zu können, was bei der Entwicklung der KI-Services und eines User Interfaces zu berücksichtigen ist.

Die Bewertung der KI-Bewertung trifft ggfs. auf andere Risiko-Bewertungen (z. B. Risikoklassen des Medizinprodukte-Durchführungsgesetz/MPDG).

3.1.3 Prüflabel für vertrauenswürdige KI

Das Prüflabel VDE SPEC 90012 V1.0 ist hilfreich, vertrauenswürdige KI-Anwendungen zu erstellen und zu bewerten.

Die in diesem Label adressierten Prüfkriterien geben bereits während der Entwicklungsphase einen Pfad vor, der in einer ethisch korrekten Anwendung mündet.

Für Anwender:innen gibt das aus anderen Prüfverfahren bekannte Label eine Orientierung über den Grad der Vertrauenswürdigkeit.

¹⁵ <https://digital-strategy.ec.europa.eu/de/policies/regulatory-framework-ai>



Abbildung 5: Kriterien der VDE SPEC 90012 V 1.0 (en); Übersetzung GSW in Abstimmung mit VDE; Quelle: Vertrauenswürdige und ethisch vertretbare KI als Grundlage für deren Akzeptanz im Rahmen des Forschungsprojektes ForeSight

3.1.4 Diskriminierungsfreie Daten als Voraussetzung vertrauenswürdiger KI

Offen zugängliche Datenquellen sind ggfs. von nicht nachweisbarer Qualität was die Ausgewogenheit ihrer Daten und der durch sie trainierten KI-Services angeht. Ohne verzerrungsfreie Daten können keine verzerrungsfreien KI-Services trainiert werden.

In den zugrundeliegenden Daten können sich Diskriminierungen zudem überlagern und verstärken und konkret das Ergebnis von KI-Anwendungen für bestimmte Gruppen von Anwender:innen verschlechtern: so werden „People of Color“ in Ki-basierten Gesichtserkennungen schlechter erkannt, sind sie außerdem noch weiblich, verschlechtert sich das Ergebnis zusätzlich.

Bsp. Im Forschungsprojekt ForeSight wurde eine KI-gestützte Gesichtserkennung als Basisdienst für einen schlüssellosen Türzugang entwickelt. Der zunächst genutzte Datensatz für die Gesichtserkennung konnte nicht sicherstellen, dass die Vielfalt möglicher Anwender:innen (28 Nationalitäten in einem Wohnprojekt) gleichermaßen Zugang zu den Türen gewährt wurde.



Innerhalb des Forschungsprojektes wurde eine Methodik für die Messung der Ausgewogenheit von Datensätzen entwickelt und ein Verfahren zur Nachbesserung. Das Ergebnis waren für diese Anwendung diskriminierungsfreie Daten und damit eine entsprechend diskriminierungsfreie Gesichtserkennung.

3.1.5 Transparenzpflicht als Voraussetzung vertrauenswürdiger KI

Ein für Anwender:innen wichtiges Feature, das während der Entwicklung berücksichtigt werden muss, ist die Transparenzpflicht.

- Diese bedeutet grundsätzlich, dass Resultate von KI-Anwendungen gegenüber Anwender:innen diese Ergebnisse erklären können müssen (siehe Kapitel 3.2.1). So darf ein Zugang per Gesichtserkennung nicht kommentarlos verboten, die Anfrage eines Kredits nicht ohne Erklärung durch eine KI-basierte Anwendung abschlägig beschieden werden.
- Weiterhin müssen durch KI-Anwendungen erzeugte oder bearbeitete Inhalte (Audios, Bilder, Videos) eindeutig als solche gekennzeichnet werden. Gelöst werden soll dies über eine Art „digitales Wasserzeichen“, das automatisch generiert und in durch KI erstellte Dateien integriert wird und nicht veränderbar ist.

3.2 Erstellung diskriminierungsfreier KI-Anwendungen

Neben der Vertrauenswürdigkeit der genutzten KI sind für diskriminierungsfreien KI-Anwendungen weitere Designmerkmale zu beachten.

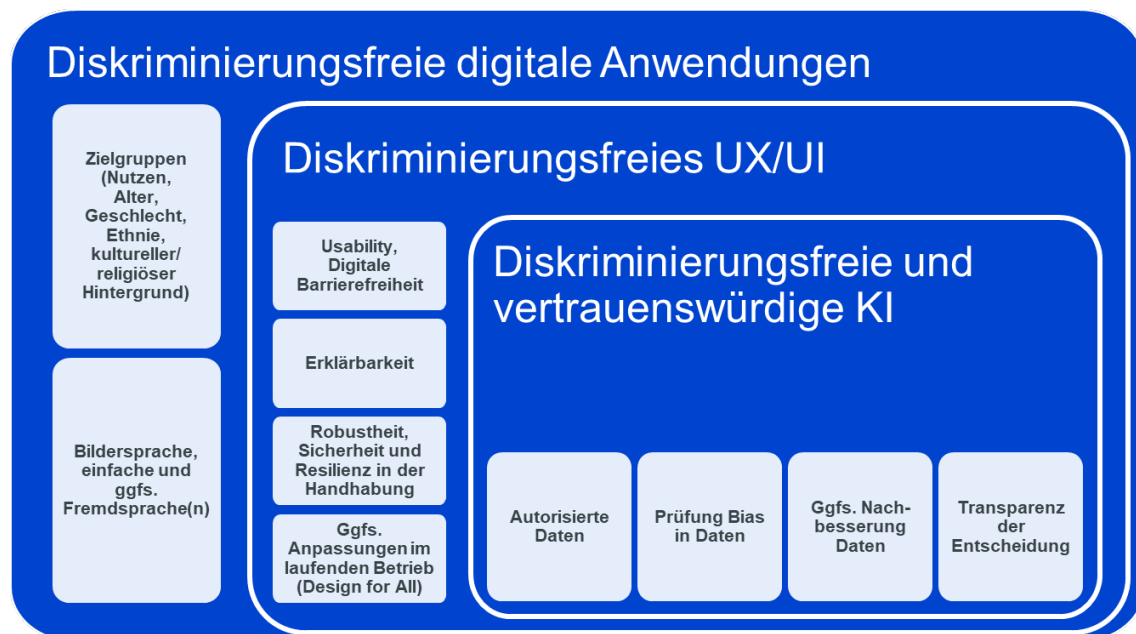


Abbildung 6: Diskriminierungsreihe KI-Anwendungen

3.2.1 UX/UI

Im Rahmen von KI-Anwendungen sind die gleichen Vorkehrungen für die Nutzbarkeit zu treffen, die generell im Rahmen barrierefreier Anwendungen und Digitalanwendungen für eine diverse Gesellschaft zu gewährleisten sind. Geregelt

wird dies durch den europäischen Accessibility ACT¹⁶ und das deutsche Barrierefreiheitsstärkungsgesetz¹⁷. Sie legen fest, dass die Nutzbarkeit durch eine Benutzerschnittstelle Anwender:innen bzw. Bevölkerungsgruppen nicht diskriminieren darf.

- Der Europäische **Rechtsakt zur Barrierefreiheit (European Accessibility Act, EAA, 2019)** soll den gleichberechtigten Zugang zu digitalen Produkten und Dienstleistungen in ganz Europa gewährleisten. Er gilt für alle Unternehmen, die in der EU tätig werden wollen, unabhängig davon, wo sie ihren Sitz haben. Sie deckt ein breites Spektrum an Waren und Dienstleistungen ab. Sie wird im Juni 2025 in allen EU-Mitgliedstaaten Gesetz.
- Mit dem **Barrierefreiheitsstärkungsgesetz (BFSG)** wurde im Juli 2021 der European Accessibility Act (EAA) ins nationale Recht überführt. Die Verordnung zum Barrierefreiheitsstärkungsgesetz (BFSGV) wurde am 15.06.2022 verabschiedet und definiert **Barrierefreiheitsanforderungen für Produkte und Dienstleistungen, die nach dem 28.06.2025 in den Verkehr gebracht bzw. für Verbraucherinnen und Verbraucher erbracht werden**. Dies umfasst u.a. den gesamten Online-Handel, Hardware, Software, aber auch überregionalen Personenverkehr oder Bankdienstleistungen.

Das Barrierefreiheitsstärkungsgesetz betrifft **Hersteller, Händler und Importeure** von bestimmten Produkten sowie Dienstleistungserbringer. Kleinunternehmen (mit weniger als 10 Beschäftigten oder weniger als 2 Mio. € Jahresumsatz) sind vom Gesetz teilweise ausgenommen. Bei **Nichteinhaltung** können Marktüberwachungsbehörden die **Bereitstellung** des Produkts oder der Dienstleistung **einschränken oder untersagen oder dafür sorgen, dass Produkte zurückgenommen oder zurückgerufen werden**. Dies betrifft nicht nur Hersteller, sondern auch Händler und Importeure.

Bsp. Die im Forschungsprojekt ForeSight in Future Living® Berlin getestete KI-gestützte Gesichtserkennung als Basisdienst für einen schlüssellosen Türzugang nutzte ein an der Wand befestigtes Tablet für die Einrichtung und Bedienung der Türöffnung durch die KI-Gesichtserkennung. Future Living® Berlin ist ein barrierefreies und in weiten Teilen rollstuhlgerechtes Wohnprojekt. Wie hoch sollte/musste das benötigte Tablet angebracht werden? Wie konnte sichergestellt werden, dass Bewohner:innen der 28 Nationalitäten nicht an der Sprachbarriere der Anwendung scheiterten? Wie sollten Kontraste gewählt und konnten Störungen durch Spiegelungen im Display vermieden werden? Wie Robust musste ein KI-System im Alltag unter Testbedingungen sein, um die Akzeptanz wirklich messen zu können? Wie sollte die Erklärbarkeit realisiert werden, wenn die KI-Anwendung zugelassene Bewohner:innen ausschloß?

¹⁶ <https://ec.europa.eu/social/main.jsp?catId=1202>, <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32019L0882>

¹⁷ <https://www.bmas.de/DE/Service/Gesetze-und-Gesetzesvorhaben/barrierefreiheitsstaerkungsgesetz.html>, <https://www.gesetze-im-internet.de/bgg/BJNR146800002.html>

Gab es unter der Bewohnerschaft Zielgruppen, für die Teilnahme am Test nicht möglich war (Kinder!)?

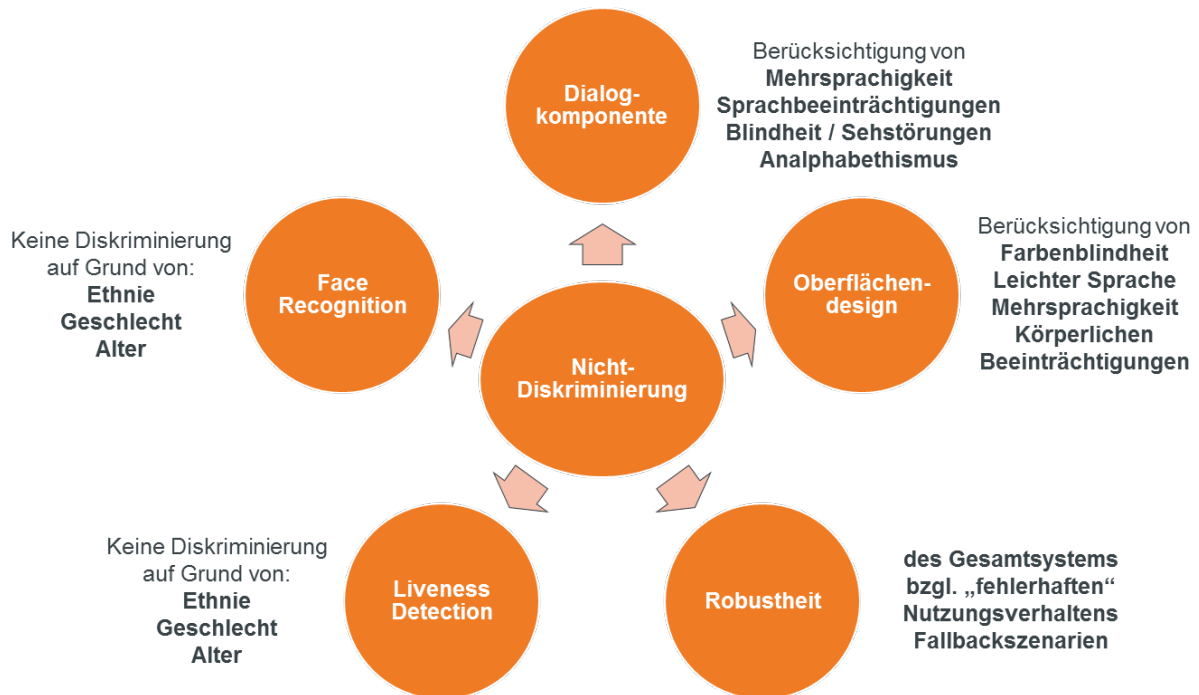


Abbildung 7: Diskriminierungsbereiche am Beispiel des KI-gestützten Gebäudepförtners in ForeSight

Die Notwendigkeit der Unterstützung digitaler Barrierefreiheit ergibt sich aus den oben genannten Gesetzen und den folgenden Zahlen:

- In Deutschland leben ca. 7,8 Mio. schwerbehinderte Menschen (Juni 2022, ca. 9,3%)¹⁸.
- Hinzu kommen Menschen mit Einschränkungen ohne Behindertenstatus:
 - Rot-grün-Schwäche (10% Männer, 0,5% Frauen)
 - Menschen mit Pflegebedarf: 5 Mio.
 - Menschen höherer Altersklassen mit schleichenden Veränderungen: 18,7 Mio. / 22,1% der Bevölkerung
 - Hören (beginnend mit höheren Tonlagen)
 - Sehen (Eintrübung der Linse, Verschlechterung Kontrastsehen)
 - Sprechen
 - Mobilität
 - Chronische Erkrankungen (Rheuma, Parkinson, MS, Krebs, ...)

¹⁸ https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Behinderte-Menschen/_inhalt.html

- Neben dauerhaften Einschränkungen durch Behinderungen gibt es temporäre (Verletzungen) oder situative Einschränkungen (Kind auf dem Arm), bei welchen das Gesetz ebenso effektiv sein kann.
- ↪ Die Umsetzung des europäischen Accessibility ACTs und des deutschen Barrierefreiheitsstärkungsgesetzes durch eine gut nutzbare, robuste und intuitiv bedienbare Anwendung trägt der Realität der Diversität der Gesellschaft Rechnung und hilft der Nutzbarkeit und vor allem der Akzeptanz jener Produkte, die es aus der Forschung hinaus in den Routinealltag schaffen, durch große Bevölkerungsgruppen.

3.3 Materialien für die Gestaltung sozialverantwortlicher KI

3.3.1 Kursreihe zur Gestaltung „Sozialverantwortlicher KI“

Der KI-Campus in Heilbronn (<https://ki-campus.org/>) bietet neben vielen anderen Online-Angeboten über das Lab für Sozialinformatik einen kostenlosen Kurs mit drei Modulen an, das KI-Entwicklung fördern will, die „allen Menschen dient, die Vielfalt und Gleichheit fördert, Diskriminierungen vermeidet“ (<https://ki-campus.org/kursreihe/ki-sozialverantwortlich-gestalten>). Er erfordert keine tieferen Kenntnisse zu KI.

Das 1. Modul „Sozialverantwortliche KI-Gestaltung“ befasst sich mit den Schwerpunkten und „Produkt“, „Gesellschaft“, „Personen“ und „Prozesse“, Viele im Leitfaden adressierten Themen werden hier aufgegriffen und vertieft werden. Er orientiert sich an dem in Heilbronn entwickelten Rahmenmodell der sozialverantwortlichen KI-Gestaltung, einem „Konzept, das darauf abzielt, Ethik und Gleichheit in den Entwicklungsprozess von KI-Technologien zu integrieren“.

Das 2. Modul „Teampraktiken für sozialverantwortliche KI-Gestaltung“ legt seinen Schwerpunkt auf die Methoden, die darin unterstützen, die Teamarbeit so zu verbessern, das in der Folge die Qualität von KI-basierten Produkten verbessert wird. Die aufgezeigten Praktiken beschränken sich nicht auf die Optimierung von Teams, deren Produkte KI nutzen. Sie lassen sich auf alle Teams übertragen.

Das 3. Modul „Methoden der Bias-Reduktion für eine sozialverantwortliche KI-Gestaltung“ legt seinen Schwerpunkt auf Praktiken der Bias-Erkennung und Reduktion in der Entwicklung und Nutzung von KI. Es will befähigen, stereotypische Verzerrungen und diskriminierende Inhalte zu identifizieren und zu vermeiden.

Alle Module sind in einzelne „Wissensnudgets“ aufgeteilt, die mit benoteten Aufgaben enden. Bei erfolgreicher Teilnahme (Erreichen einer Punktzahl) kann eine Teilnahmebestätigung oder ein Leistungsnachweis erstellt werden.

3.3.2 „Faires KI-Prompting - Ein Leitfaden für Unternehmen“

Die BSP Business and Law School – Hochschule für Management und Recht, Projekt Mittelstand-Digital Zentrum Zukunftskultur, hat einen Leitfaden

veröffentlicht¹⁹, der Nutzenden von KI-Anwendungen ein Werkzeug an die Hand geben will, um vor allem Anwendungen aus dem Bereich generativer KI verantwortungsvoll und bewusst zu nutzen. Diese Art der KI-Anwendungen basieren auf trainierten Modellen und Algorithmen. Sie werden durch Eingaben, sogenannte „Prompts“, abgefragt und sind in der Lage, Inhalte wie z.B. Texte, Bilder, Musik, Videos oder sogar Code zu erzeugen.

Der Leitfaden will Verständnis dafür schaffen, welche positiven wie auch negativen Ergebnisse durch die Eingaben von der Nutzung generativer KI geschaffen werden können und zugleich das Verständnis und den Weg für einen diversen und fairen Einsatz ebnen.

¹⁹ Gengler, E., Kraus, A., Bodrožić-Brnić, K. (2024). Faires KI-Prompting – Ein Leitfaden für Unternehmen. BSP Business and Law School – Hochschule für Management und Recht. (1 – 28)
<https://www.digitalzentrum-zukunftskultur.de/wp-content/uploads/2024/05/Faires-KI-Prompting-Ein-Leitfaden-fuer-Unternehmen.pdf>