

KI und UX mit Impact gestalten: Nutzer*innenbedürfnisse, Herausforderungen und Handlungsempfehlungen für ein KI-gestütztes Suchtool am Beispiel der Globetrotter-Website

BACHELOR THESIS

Juni 2025

Autor
Michael Gross

Betreuungsperson
Nico Erb

Praxispartner*in
campfire GmbH
Pius Caduff

Abstract

Diese Bachelorarbeit untersucht die spezifischen Nutzer*innenbedürfnisse sowie Herausforderungen bei der Nutzung eines KI-gestützten Suchtools auf der Globetrotter-Website. Daraus werden praxisnahe Handlungsempfehlungen zur Optimierung der User Experience abgeleitet. Für die Analyse wurden qualitative Remote-Nutzer*innentests durchgeführt ($N = 7$). Die Testpersonen waren zwischen 23 und 60 Jahre alt und verfügten über unterschiedliche Erfahrungen mit Künstlicher Intelligenz und Reiseagenturen. Zur Datenerhebung kamen die Think-Aloud-Methode sowie eine leitfadengestützte Vor- und Nachbefragung zum Einsatz. Als theoretisches Fundament dienten die ISO 9241-210 für menschenzentrierte Gestaltung, das UX-Honeycomb-Modell zur umfassenden UX-Evaluation und das Technologieakzeptanzmodell zur Analyse der Akzeptanzfaktoren. Die Ergebnisse zeigen, dass Nutzende eine effiziente, intuitive und kontextsensitive Informationssuche sowie eine interaktive und transparente Kommunikation erwarten. Als relevant wurden ausserdem Aspekte wie Glaubwürdigkeit, Aktualität und Nachvollziehbarkeit der KI-Ergebnisse genannt. Herausforderungen bestehen unter anderem in der teilweise unklaren Auffindbarkeit beim Gebrauch des KI-Suchtools, unpräzisen oder zu knappen Antworten, technischen Problemen wie langen Ladezeiten sowie ablenkenden oder überladenen Interface-Designs. Daraus abgeleitete Empfehlungen umfassen eine deutlichere Sichtbarkeit und Kennzeichnung des KI-Suchtools, transparente Erwartungssteuerung hinsichtlich Leistungsgrenzen, technische Optimierungen zur Steigerung der Performance sowie eine Weiterentwicklung der KI-Technologie für präzisere und umfassendere Antworten. Zudem wird empfohlen, dass das KI-Suchtool vor dem Antworten kontextbezogene Rückfragen stellt, um die Nutzer*innenanliegen besser zu verstehen. Ergänzt werden diese Massnahmen durch Vorschläge für eine übersichtliche, ablenkungsfreie Gestaltung. Diese Arbeit unterstreicht, dass für eine nachhaltige Verbesserung der User Experience eines KI-Suchtools zentrale UX-Prinzipien und eine iterative, menschenzentrierte Weiterentwicklung essenziell sind.

Anzahl Zeichen: 124'429

Anzahl Worte Abstract: 232

Keywords: Künstliche Intelligenz, KI-Suchtool, Menschenzentrierte Gestaltung, User Experience, Nutzer*innentest

Inhaltsverzeichnis

1	Einleitung	1
1.1	Ausgangslage und Problemstellung.....	1
1.2	Zielsetzung und Fragestellung.....	2
	Erhebung der Nutzer*innenbedürfnisse.	2
	Evaluierung des KI-Kompasses.	2
	Ableitung von Handlungsempfehlungen.....	2
1.3	Aufbau der Arbeit.....	3
2	Theoretische Grundlagen	4
2.1	Menschenzentrierte Gestaltung	4
2.1.1	Grundlagen und Prinzipien der menschenzentrierten Gestaltung	4
2.1.2	Iterativer Prozess der menschenzentrierten Gestaltung.....	5
	Analyse des Nutzungskontexts.	5
	Spezifikation der Nutzungsanforderungen.	5
	Gestaltung von Lösungen.	5
	Evaluation durch Nutzer*innentests.	6
	Implementierung und kontinuierliche Verbesserung.....	6
2.2	Technologieakzeptanzmodell	6
2.3	User Experience	8
2.3.1	UX-Honeycomb-Modell.....	8
	Nützlichkeit.	9
	Benutzbarkeit.....	9
	Attraktivität.....	10
	Auffindbarkeit.....	10
	Zugänglichkeit.....	10
	Glaubwürdigkeit.....	10
	Wertstiftung.....	11
3	Methodik.....	12
3.1	Nutzer*innentests	12

UX-Evaluation: Globetrotter KI-Kompass

3.1.1	Think-Aloud-Methode	13
3.1.2	Pretest.....	13
3.1.3	Leitfadenerstellung Vor- und Nachbefragung.....	13
3.1.4	Testaufgaben.....	15
3.2	Sampling	16
3.3	Datenauswertung	17
4	Ergebnisse	20
4.1	Erfahrungen und Einstellungen.....	20
4.1.1	Erfahrungen und Einstellungen zu KI-Tools.....	20
4.1.2	Herausforderungen bei klassischen Suchfunktionen.....	20
4.2	Mentale Modelle und Erwartungen	21
4.2.1	Erwartungen an eine Suchfunktion	21
4.2.2	Erwartungen an den KI-Kompass	21
4.2.3	Suchstrategien.....	22
4.3	Nutzer*innenbedürfnisse	22
4.3.1	Effiziente Informationsfindung.....	22
4.3.2	Kontextsensitive und Interaktive Kommunikation	23
4.3.3	Intuitive und übersichtliche Gestaltung.....	23
4.3.4	Glaubwürdige und aktuelle Ergebnisse.....	24
4.4	Interaktion mit dem KI-Kompass.....	24
4.4.1	Erste Reaktion auf den KI-Kompass	24
4.4.2	Suchverhalten mit dem KI-Kompass.....	25
4.4.3	Umgang mit Links	25
4.5	Evaluation des KI-Kompasses anhand des UX-Honeycomb-Modells	26
4.5.1	Nützlichkeit	27
4.5.2	Benutzbarkeit.....	28
4.5.3	Attraktivität.....	28
4.5.4	Auffindbarkeit.....	30
4.5.5	Zugänglichkeit	31

4.5.6	Glaubwürdigkeit.....	31
4.5.7	Wertstiftung	33
4.6	Identifizierte Herausforderungen.....	34
4.7	Identifizierte Optimierungspotenziale	37
5	Diskussion.....	39
5.1	Beantwortung der Fragestellung.....	39
5.1.1	Nutzer*innenbedürfnisse.....	39
	Effiziente Informationsfindung.....	39
	Kontextsensitive und interaktive Kommunikation.....	40
	Intuitive und übersichtliche Gestaltung.....	40
	Glaubwürdige und aktuelle Ergebnisse.....	41
5.1.2	Identifizierte Herausforderungen.....	41
	Schlechte Auffindbarkeit.....	41
	Unpräzise Antworten.....	42
	Technische Probleme und Ablenkungen.....	42
5.1.3	Handlungsempfehlungen.....	42
	Auffindbarkeit und Erwartungssteuerung verbessern.....	42
	Dialogorientierung und Kontextualisierung stärken.....	44
	Inhaltsqualität und Mehrwert ausbauen.....	44
	Stabilität und Leistung verbessern.....	44
	Nutzer*innenzentrierte Weiterentwicklung.....	44
	Zusammenfassende Beantwortung der Forschungsfrage.....	45
6	Fazit	46
6.1	Limitationen und Reflexion des Forschungsdesigns	46
6.2	Ausblick und Forschungsimplicationen	47
6.3	Schlusswort.....	48
	Literaturverzeichnis	49
	Abbildungsverzeichnis.....	52
	Abkürzungsverzeichnis.....	54

Hilfsmittelverzeichnis mit Verwendungszweck55

1 Einleitung

Die Nutzung digitaler Produkte und Dienstleistungen ist in den letzten Jahren stark gestiegen, da immer mehr Menschen täglich auf Webseiten, Apps und andere digitale Angebote zugreifen (Moser, 2012). Dabei spielt die User Experience (UX) eine entscheidende Rolle für den Erfolg dieser digitalen Produkte. Eine positive UX fördert die Zufriedenheit und Bindung, während eine unzureichende UX oft dazu führt, dass Nutzende abspringen und Alternativen suchen (Moser, 2012).

Vor diesem Hintergrund beschäftigt sich diese Arbeit mit der Frage, wie ein mit künstlicher Intelligenz (KI) gestütztes Suchtool gestaltet sein muss, um den tatsächlichen Bedürfnissen der Nutzenden gerecht zu werden und die UX nachhaltig zu optimieren.

In dieser Arbeit wird *Suchfunktion* als übergeordneter Begriff für sämtliche Suchsysteme verwendet. *Klassische Suchfunktion* bezeichnet herkömmliche, nicht-KI-basierte Systeme, während für das KI-gestützte System der Begriff *KI-Suchtool* genutzt wird. Der Begriff *Benutzer*innenfreundlichkeit* (engl. Usability) wird in dieser Arbeit entsprechend der ISO 9241-210 (2020) als *Gebrauchstauglichkeit* definiert.

Die folgenden Abschnitte dieses Kapitels skizzieren zunächst die Problemstellung und den Kontext des Themas, bevor die Zielsetzung der Untersuchung, die zugrunde liegende Forschungsfrage und der Ablauf der Arbeit vorgestellt werden.

1.1 Ausgangslage und Problemstellung

In der heutigen digitalen Landschaft ist die effiziente und zielgerichtete Informationssuche ein zentraler Bestandteil der UX von Webseiten und Anwendungen. Gerade im E-Commerce und Dienstleistungsbereich ist die Suchfunktion zentral für die Nutzer*innenzufriedenheit (Moser, 2012). Klassische Suchfunktionen stossen jedoch häufig an ihre Grenzen und erfüllen die Erwartungen der Nutzenden nicht immer optimal (Spatharioti, Rothschild, Goldstein & Hofman, 2023). Mit dem Aufkommen von KI eröffnen sich neue Möglichkeiten, die Such- und Navigationserlebnisse grundlegend zu verbessern (Xiong et al., 2024).

Die campfire GmbH arbeitet an der Entwicklung eines KI-Suchtools, das die Bedürfnisse der Nutzenden gezielt adressieren soll. Die Idee ist, neben der klassischen Suchfunktion, eine alternative Möglichkeit zu bieten, um Inhalte zu durchsuchen und Antworten auf offene Fragen zu finden. Dadurch sollen Nutzende besser unterstützt werden, um ihre individuellen Ziele im jeweiligen Nutzungskontext adäquat zu erreichen. Seit Mitte Januar ist die Beta-Version des KI-Suchtools namens *KI-Kompass* auf der Website von Globetrotter online. Die Integration des KI-Suchtools soll für campfire und Globetrotter die UX verbessern, die Effizienz und Kundenzufriedenheit steigern, die Marktposition stärken und das Markenimage innovativ prägen. Darüber hinaus ermöglicht

das KI-Suchtool die gezielte Erhebung von Nutzer*innendaten, um Optimierungspotenziale zu identifizieren.

Die zentrale Herausforderung besteht darin, ein KI-Suchtool so zu gestalten und zu integrieren, dass es die Erwartungen und Anforderungen der Nutzenden erfüllt. Ein häufiges Problem bei der Einführung neuer Technologien liegt darin, dass diese zwar technisch leistungsfähig, jedoch nicht menschenzentriert entwickelt sind (Moser, 2012). Wenn die Bedürfnisse und mentalen Modelle der Nutzenden nicht ausreichend berücksichtigt werden, führt dies oft zu einer niedrigen Akzeptanz und begrenztem Mehrwert für die Nutzenden (Davis, 1989).

Bisher fehlen Erkenntnisse darüber, wie ein solches KI-Suchtool optimal gestaltet werden soll, um die UX tatsächlich zu verbessern. Insbesondere ist unklar, welche spezifischen Anforderungen und Erwartungen die Nutzenden an ein KI-Suchtool haben.

1.2 Zielsetzung und Fragestellung

Das übergeordnete Ziel ist die Evaluation des bestehenden KI-Suchtools auf der Globetrotter-Website, um Nutzer*innenbedürfnisse und Optimierungsmöglichkeiten zu identifizieren, damit es für Nutzende einen Mehrwert bietet.

Dabei sollen folgende spezifische Ziele verfolgt werden:

Erhebung der Nutzer*innenbedürfnisse. Durch Nutzer*innentests inklusive Vor- und Nachbesprechung sollen die Anforderungen, Bedürfnisse und mentalen Modelle der Zielgruppe identifiziert werden. Diese Erkenntnisse sind essenziell, da nur auf Basis eines fundierten Verständnisses der Nutzer*innenperspektive ein System gestaltet werden kann, das als hilfreich, akzeptiert und intuitiv wahrgenommen wird (ISO 9241-210, 2020). Die Analyse von Nutzungskontext, Einstellung zu KI, mentalen Modellen und Herausforderungen mit Suchfunktionen bildet die Basis für die gezielte Optimierung des KI-Suchtools.

Evaluierung des KI-Kompasses. Die aktuelle Version des KI-Suchtools wird anhand der Globetrotter-Website systematisch analysiert, um Herausforderungen zu identifizieren. Diese Erkenntnisse sind essenziell, um UX-Probleme gezielt zu beheben und das KI-Suchtool bedarfsgerecht weiterzuentwickeln (ISO 9241-210, 2020).

Ableitung von Handlungsempfehlungen. Die gewonnenen Erkenntnisse führen zu praxisnahen Handlungsempfehlungen, die sowohl für Globetrotter als auch für künftige Integrationen auf anderen Websites relevant sind. Diese Empfehlungen dienen als strategische Grundlage für eine menschenzentrierte Weiterentwicklung des KI-Suchtools. Im Ausblick dieser Arbeit werden daraus Handlungsempfehlungen abgeleitet, wie KI-Suchtools allgemein gestaltet sein müssen, um langfristig eine positive, vertrauenswürdige und effektive UX zu ermöglichen.

Auf Basis dieser Ziele wurde folgende Fragestellung abgeleitet:

Welche spezifischen Nutzer*innenbedürfnisse bestehen im Zusammenhang mit einem KI-gestützten Suchtool auf der Globetrotter-Website, welche Herausforderungen ergeben sich bei dessen Nutzung und welche konkreten Handlungsempfehlungen lassen sich daraus ableiten, um die User Experience gezielt zu verbessern?

Diese Arbeit soll damit einen Beitrag zur Gestaltung menschenzentrierter KI-Suchtools leisten und praxisrelevante Erkenntnisse für die UX-Optimierung von digitalen Lösungen bereitstellen.

1.3 Aufbau der Arbeit

Diese Arbeit ist in mehrere aufeinander aufbauende Kapitel gegliedert. Nach der Einleitung, in der die Ausgangslage, die Problemstellung sowie die Zielsetzung und Forschungsfrage erläutert wurden, folgt das Kapitel zu den theoretischen Grundlagen. In diesem Abschnitt werden zentrale Konzepte und Modelle der UX vorgestellt, die für die Untersuchung relevant sind, darunter die menschenzentrierte Gestaltung nach ISO 9241-210, das UX-Honeycomb-Modell von Peter Morville und das Technologieakzeptanzmodell (TAM) von Fred Davis. Diese theoretische Fundierung bildet die Basis für die methodische Herangehensweise der Arbeit.

Im dritten Kapitel wird die Methodik der Untersuchung beschrieben. Hier wird das Forschungsdesign, die Auswahl der Stichprobe sowie die Konzipierung, Durchführung und Auswertung der Nutzer*innentests erläutert. Der Fokus liegt dabei auf qualitativen Methoden, die es ermöglichen, tiefgehende Erkenntnisse über die Nutzer*innenbedürfnisse und Herausforderungen bei der Nutzung des KI-Suchtools zu gewinnen.

Das vierte Kapitel präsentiert die Ergebnisse der Untersuchung. Es umfasst die Analyse der Nutzer*innenbedürfnisse, die Evaluation des bestehenden Prototyps sowie die Identifikation von Herausforderungen und Optimierungspotenzialen. Diese Erkenntnisse bilden die Grundlage für die anschließende Diskussion.

In Kapitel fünf werden die gewonnenen Ergebnisse im Hinblick auf die Forschungsfrage diskutiert. Dabei erfolgt zunächst die Beantwortung der Forschungsfrage anhand der identifizierten Nutzer*innenbedürfnisse und Herausforderungen und den daraus abgeleiteten Handlungsempfehlungen.

Das sechste und abschliessende Kapitel bildet das Fazit der Arbeit. Es werden die Limitationen der Untersuchung aufgezeigt und das Forschungsdesign kritisch reflektiert. Im Ausblick werden die Implikationen der Untersuchungsergebnisse für die zukünftige Gestaltung und Implementierung menschenzentrierter KI-Suchtools auf anderen Websites diskutiert.

2 Theoretische Grundlagen

Für eine fundierte Untersuchung des KI-Suchtools ist es notwendig, zentrale theoretische Konzepte der menschenzentrierten Gestaltung, der Technologieakzeptanz und der UX zu berücksichtigen. Die folgenden Abschnitte liefern die konzeptionelle Basis für die spätere Analyse und die Ableitung von Handlungsempfehlungen. Neben der menschenzentrierten Gestaltung nach ISO 9241-210 werden auch praxisrelevante Modelle wie das TAM sowie das UX-Honeycomb-Modell vorgestellt. Diese theoretischen Ansätze helfen dabei, sowohl die Bedürfnisse der Nutzenden als auch die Herausforderungen in der Nutzung des KI-Suchtools besser zu verstehen und zu bewerten.

2.1 Menschenzentrierte Gestaltung

Die menschenzentrierte Gestaltung ist ein zentrales Konzept in der Entwicklung interaktiver Systeme und bildet das methodische Fundament dieser Arbeit. Sie stellt sicher, dass Produkte und Dienstleistungen systematisch auf die Bedürfnisse, Fähigkeiten und Erwartungen der Nutzenden ausgerichtet werden. Grundlage ist die ISO 9241-210 (2020), welche menschenzentrierte Gestaltung als iterativen Prozess definiert, in dem Nutzende aktiv in den gesamten Entwicklungszyklus eingebunden werden. Ziel ist die Entwicklung gebrauchstauglicher, effektiver und zufriedenstellender Systeme.

2.1.1 Grundlagen und Prinzipien der menschenzentrierten Gestaltung

Menschenzentrierte Gestaltung geht über die reine technische Funktionalität hinaus und betrachtet das System konsequent aus der Perspektive der Nutzenden. Der Prozess beginnt mit einer umfassenden Analyse des Nutzungskontextes. Die relevanten Nutzer*innengruppen, deren Aufgaben sowie die physische, soziale und technische Umgebung werden detailliert untersucht. Dabei wird betont, dass sowohl funktionale als auch emotionale, soziale und kognitive Aspekte der Nutzung berücksichtigt werden müssen (ISO 9241-210, 2020).

Ein zentrales Prinzip ist die kontinuierliche und aktive Einbeziehung der Nutzenden im Entwicklungsprozess. Durch regelmässige Rückmeldungen, Nutzer*innentests und iterativen Optimierungen wird gewährleistet, dass die entwickelten Lösungen den tatsächlichen Anforderungen entsprechen und die Gebrauchstauglichkeit kontinuierlich verbessert wird. Entscheidungen im Gestaltungsprozess sollten dabei empirisch fundiert sein, zum Beispiel auf Basis von Nutzer*innentests oder Beobachtungen (ISO 9241-210, 2020).

Eine nutzer*innenzentrierte Entwicklung erfordert zudem eine interdisziplinäre Zusammenarbeit, insbesondere zwischen UX-Design, Psychologie, Informatik und weiteren relevanten Disziplinen. Nur so kann eine ganzheitliche, positive UX entstehen, die Akzeptanz, Effizienz und Zufriedenheit der Nutzenden fördert (ISO 9241-210, 2020).

2.1.2 Iterativer Prozess der menschenzentrierten Gestaltung

Die menschenzentrierte Gestaltung verläuft nach ISO 9241-210 (2020) in einem iterativen Prozess, der aus mehreren Phasen besteht (vgl. Abb. 1).

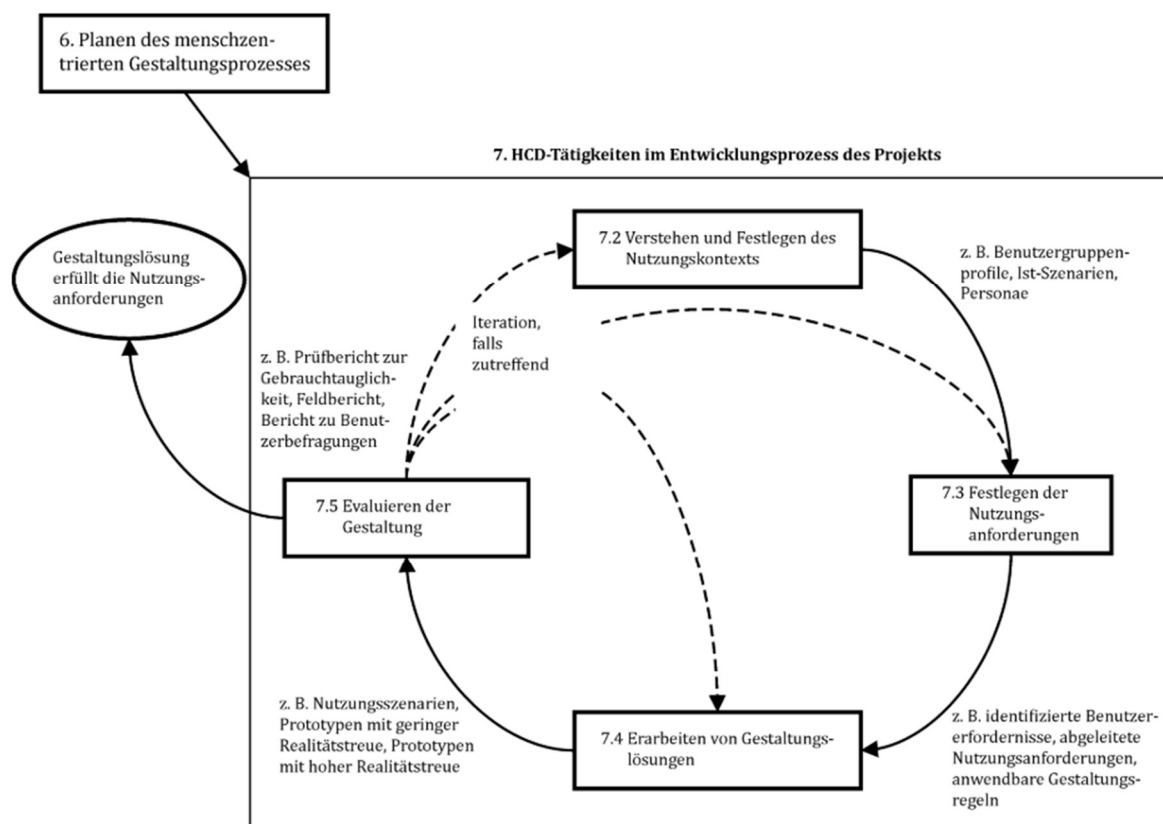


Abbildung 1. Die Phasen der menschenzentrierten Gestaltung, von der Kontextanalyse bis zur Evaluation, sind in einem wechselseitig abhängigen und iterativen Zyklus organisiert (ISO 9241-210, 2020, S. 21)

Analyse des Nutzungskontexts. Die erste Phase beinhaltet die Identifikation der relevanten Nutzer*innengruppen und eine detaillierte Analyse ihrer Aufgaben, Ziele und Nutzungsszenarien. Es werden Methoden wie Interviews, Fokusgruppen und ethnografische Studien eingesetzt, um ein tiefgehendes Verständnis für die Anforderungen und Herausforderungen der Nutzenden zu gewinnen (ISO 9241-210, 2020).

Spezifikation der Nutzungsanforderungen. Basierend auf den Erkenntnissen der Analysephase werden konkrete Anforderungen an das interaktive System formuliert. Diese Anforderungen müssen sowohl funktionale als auch nicht-funktionale Aspekte der Nutzung berücksichtigen und bilden die Grundlage für das Design und die Entwicklung des Systems (ISO 9241-210, 2020).

Gestaltung von Lösungen. In dieser Phase werden erste Konzepte, Wireframes, Prototypen und Designentwürfe erstellt. Die Nutzenden werden aktiv in den Prozess einbezogen, um sicherzustellen, dass entstehende Systeme ihren Anforderungen entsprechen. Die Gestaltung erfolgt iterativ, sodass frühzeitig Optimierungen vorgenommen werden können (ISO 9241-210, 2020).

Evaluation durch Nutzer*innentests. Die entwickelten Prototypen werden in realistischen Nutzungsszenarien getestet. Dabei kommen sowohl qualitative Methoden wie Beobachtungen und Interviews als auch quantitative Messungen zur Gebrauchstauglichkeit zum Einsatz. Auf Grundlage der Evaluationsergebnisse werden anschliessend notwendige Anpassungen und Verbesserungen vorgenommen (ISO 9241-210, 2020).

Implementierung und kontinuierliche Verbesserung. Nach der finalen Entwicklung und Implementierung des Systems wird weiterhin sichergestellt, dass es kontinuierlich verbessert wird. Nutzer*innenfeedback, Fehleranalysen und technologische Entwicklungen fliessen in regelmässige Optimierungsprozesse ein (ISO 9241-210, 2020).

Die konsequente Anwendung der Prinzipien menschenzentrierter Gestaltung nach ISO 9241-210 (2020) bildet die theoretische Grundlage dieser Arbeit. Sie legitimiert die Wahl qualitativer Methoden, die aktive Einbeziehung der Nutzenden und die iterative Weiterentwicklung des KI-Suchtools. Die Orientierung an diesen Prinzipien stellt sicher, dass die abgeleiteten Handlungsempfehlungen empirisch fundiert und wissenschaftlich abgesichert sind. Die ISO 9241-210 diente als Grundlage zur Formulierung der Hypothesen für die Nutzer*innentests sowie für die Entwicklung des Interviewleitfadens und ist entsprechend auch im Kategoriensystem der Auswertung verankert.

2.2 Technologieakzeptanzmodell

Das von Fred Davis entwickelte TAM ist ein zentrales Modell zur Untersuchung der Faktoren, die die Nutzung und Akzeptanz von Technologien beeinflussen (Davis, 1985; Marangunić & Granić, 2015).

Studien zeigen, dass die Absicht zur Nutzung der stärkste Prädiktor für die tatsächliche Nutzung eines Systems ist. Diese Nutzungsabsicht wird massgeblich von der Einstellung gegenüber der Systemnutzung beeinflusst. Die Einstellung wiederum basiert auf zwei zentralen Faktoren, dem wahrgenommenen Nutzen und der wahrgenommenen Einfachheit der Bedienung (Davis & Venkatesh, 1996).

Der wahrgenommene Nutzen beschreibt die subjektive Einschätzung einer Person, inwieweit die Nutzung eines bestimmten Systems für sie vorteilhaft ist. Ein System gilt als besonders nützlich, wenn Nutzende erwarten, dadurch einen Mehrwert oder eine spürbare Erleichterung zu erzielen (Davis, 1989; Davis & Venkatesh, 1996).

Die wahrgenommene Einfachheit der Bedienung hingegen bezieht sich darauf, wie sehr eine Person davon überzeugt ist, das System ohne grossen Aufwand nutzen zu können (Davis, 1989; Davis & Venkatesh, 1996).

Abbildung 2 bietet eine Übersicht über die zentralen Zusammenhänge des TAM.

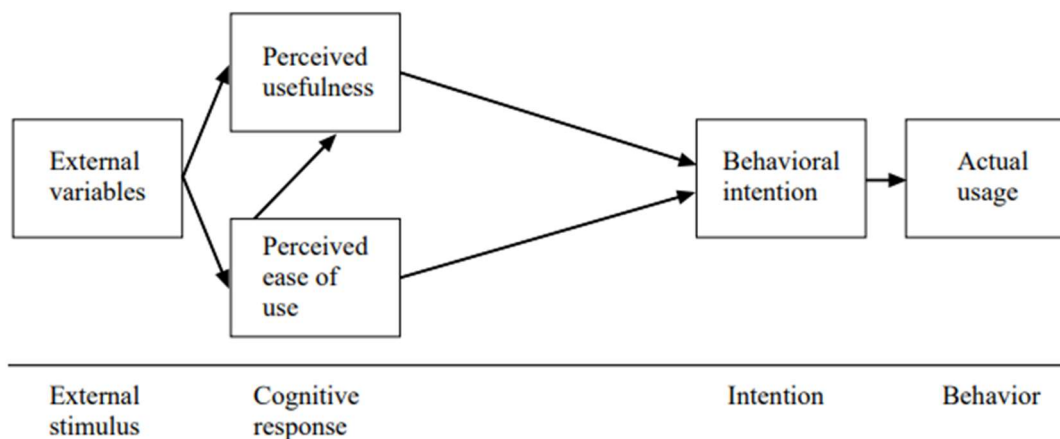


Abbildung 2. Das TAM: Die wahrgenommene Nützlichkeit und die wahrgenommene Einfachheit der Bedienung beeinflussen die Benutzungsintention und somit die tatsächliche Benutzung (Davis & Venkatesh, 1996, S. 2)

Dieses Konzept verdeutlicht, wie einfach und intuitiv eine Technologie subjektiv erscheint. Je einfacher die Bedienung eines Systems eingeschätzt wird, desto niedriger sind die mentalen und physischen Hürden der Nutzung, was wiederum die Akzeptanz fördert (Marangunić & Granić, 2015). Dies ist besonders relevant, da kognitive und physische Ressourcen der Nutzenden begrenzt sind (Davis, 1989).

Nach dem TAM beeinflussen sowohl der wahrgenommene Nutzen als auch die wahrgenommene Einfachheit der Bedienung die Einstellung gegenüber der Technologie. Diese Einstellung wirkt sich wiederum auf die Verhaltensabsicht aus, die als entscheidender Faktor für das tatsächliche Nutzungsverhalten gilt (Marangunić & Granić, 2015). Empirische Studien zeigen zudem, dass eine höhere Gebrauchstauglichkeit indirekt die Nutzungsabsicht steigern kann, da einfach zu bedienende Systeme oft auch als nützlicher wahrgenommen werden. Wenn eine Technologie weniger Aufwand erfordert, um gewünschte Ergebnisse zu erzielen, steigt deshalb ihre Attraktivität für potenzielle Nutzende (Marangunić & Granić, 2015).

Die kontinuierliche Weiterentwicklung des TAM hat dazu beigetragen, dass es heute zu den wichtigsten Modellen zur Erklärung und Prognose von Technologienutzung zählt (Davis & Venkatesh, 1996; Marangunić & Granić, 2015).

Obwohl das TAM zur Erklärung der Akzeptanz neuer Technologien entwickelt wurde, sind seine Kernkonzepte auch für die Optimierung bestehender Tools relevant. Verbesserungen am KI-Suchtool zielen darauf ab, die wahrgenommene Nützlichkeit und Einfachheit der Bedienung zu steigern und damit die Nutzung und Zufriedenheit zu erhöhen. Das TAM wurde genutzt, um die Interviewfragen gezielt auf die Erfassung mentaler Modelle und Einstellungen der Nutzenden auszurichten, Hypothesen für die Nutzer*innentests zu entwickeln und daraus passende Testaufgaben abzuleiten. Zudem

half es dabei, die Ergebnisse gezielt im Hinblick auf konkrete UX-Potenziale und Herausforderungen zu interpretieren. Das Modell liefert dadurch wertvolle Anhaltspunkte für gezielte Optimierungen aus psychologischer Perspektive.

2.3 User Experience

Die ISO 9241-210 (2020) beschreibt UX wie folgt: «Wahrnehmungen und Reaktionen einer Person, die aus der tatsächlichen und/oder der erwarteten Benutzung eines Systems, eines Produktes oder einer Dienstleistung resultieren» (S. 10).

UX bezieht sich daher nicht nur auf die unmittelbare Nutzung eines Produkts, sondern auch auf die Eindrücke und Wirkungen, die es vor und nach der eigentlichen Interaktion bei Nutzenden hinterlässt (Richter & Flückiger, 2016). Durch den Fokus auf das gesamte Nutzungserlebnis überschneidet sich das Feld zunehmend mit Bereichen wie Produktdesign, User Interface (UI)-Gestaltung und der Optimierung angrenzender Prozesse. Aufgrund dieses ganzheitlichen Ansatzes ersetzt der Begriff UX zunehmend die klassische Bezeichnung Gebrauchstauglichkeit (Richter & Flückiger, 2016).

Im folgenden Kapitel wird ein Modell präsentiert, welches zur Analyse und Bewertung der UX eines Produkts oder einer Website dient.

2.3.1 UX-Honeycomb-Modell

Das UX-Honeycomb-Modell von Peter Morville stellt eine der umfassendsten Ansätze zur Analyse und Optimierung der UX dar. Es geht über die reine Gebrauchstauglichkeit hinaus und umfasst sieben zentrale Facetten der Nutzungserfahrung. Nützlichkeit, Benutzbarkeit, Attraktivität, Auffindbarkeit, Zugänglichkeit, Glaubwürdigkeit und Wertstiftung bilden gemeinsam die Grundlage für die Gestaltung effektiver, menschenzentrierter digitaler Produkte und Dienstleistungen (Morville, 2004).

Die Stärke dieses Modells liegt in seiner ganzheitlichen Perspektive auf die UX-Gestaltung. Traditionelle UX-Ansätze konzentrierten sich vor allem auf Gebrauchstauglichkeit, während das UX-Honeycomb-Modell zeigt, dass herausragende UX ein ausgewogenes Zusammenspiel mehrerer Faktoren erfordert. (Morville, 2004).

Abbildung 3 gibt einen Überblick über die sieben Facetten des UX-Honeycomb-Modells.



Abbildung 3. Die sieben Facetten des UX-Honeycomb-Modells erinnern an eine Bienenwabe, um die enge Verzahnung der UX-Faktoren zu symbolisieren. Eigene Darstellung nach Morville (2004)

Nützlichkeit. Ein digitales Produkt oder eine Dienstleistung sollte einen klaren Nutzen für die Nutzenden bieten. Das bedeutet, dass UX-Designer sich nicht nur an bestehenden Strukturen orientieren sollten, sondern auch innovative Lösungen entwickeln müssen, die tatsächliche Probleme der Nutzenden lösen. Ein Produkt, welches gut gestaltet ist, jedoch keinen erkennbaren Nutzen hat, wird nicht erfolgreich sein (Morville, 2004). Nicht nur Morville hebt in seinem UX-Honeycomb-Modell die Bedeutung von Nützlichkeit hervor, auch andere angesehene UX-Experten betonen, dass ein Produkt nur dann eine gute Nutzungserfahrung bieten kann, wenn es den Nutzenden tatsächlich einen funktionalen Mehrwert bietet (Diefenbach & Hassenzahl, 2017; Nielsen, 1994; Norman, 2013).

Benutzbarkeit. Die Benutzbarkeit ist eine der bekanntesten und wichtigsten Dimensionen der UX. Ein System sollte intuitiv, einfach zu bedienen und effizient sein. Während Benutzbarkeit eine notwendige Voraussetzung für eine gute UX darstellt, ist sie allein jedoch nicht ausreichend. Erst in Kombination mit den anderen Facetten des UX-Honeycomb-Modells wird die Benutzbarkeit zu einem integralen Bestandteil der Gesamtnutzungserfahrung (Morville, 2004). Gleichwohl bleibt die Benutzbarkeit ein entscheidendes Qualitätskriterium. Studien belegen, dass eine schlechte Benutzbarkeit schnell zu Frustration führt und Nutzer*innenabwanderung bewirkt (Tuch, Presslauer, Stöcklin, Opwis & Bargas-Avila, 2012). Im UX-Honeycomb-Modell wird der Begriff benutzbar (engl. usable) als Synonym für gebrauchstauglich verwendet. Laut ISO 9241-210

(2020) beschreibt gebrauchstauglich die Eigenschaft eines Systems, von Nutzenden effektiv, effizient und zufriedenstellend genutzt werden zu können.

Attraktivität. Neben Nützlichkeit und Benutzbarkeit spielt auch die emotionale Komponente eine zentrale Rolle. Visuelles Design, Markenidentität und Interaktionselemente beeinflussen die Wahrnehmung eines Produkts und schaffen eine emotionale Bindung zur Marke. Attraktive, ansprechende und konsistente Designs können die Nutzungserfahrung positiv verstärken und die Markenwahrnehmung verbessern (Morville, 2004). Ähnlich spricht Hassenzahl (2004) von hedonischen Qualitäten eines interaktiven Produkts. Freude, Stimulation, ästhetische Anziehungskraft und Spass bei der Nutzung tragen dazu bei, dass ein Produkt als erlebenswert wahrgenommen wird. Empirische Forschung untermauert dies. So finden Tractinsky, Katz und Ikar (2000), dass Nutzende schön gestaltete Benutzungsoberflächen oft auch als gebrauchstauglicher einstufen.

Auffindbarkeit. Ein Produkt oder eine Website muss so gestaltet sein, dass Nutzende relevante Inhalte leicht finden können. Auffindbarkeit bezieht sich sowohl auf die Informationsarchitektur als auch auf die Navigation und Suchfunktionen innerhalb eines digitalen Angebots. In Zeiten der Suchmaschinenoptimierung spielt dieser Aspekt eine entscheidende Rolle, da Inhalte auch ausserhalb der eigenen Plattform auffindbar sein müssen (Morville, 2004). Gerade im Web-Kontext ist Auffindbarkeit kritisch. Inhalte, die nicht gefunden werden, existieren aus Sicht der Nutzenden quasi nicht (Morville, 2005).

Zugänglichkeit. Digitale Produkte sollten für alle Nutzende zugänglich sein, einschliesslich Menschen mit Behinderungen. Barrierefreiheit ist nicht nur eine ethische Verpflichtung, sondern wird zunehmend auch gesetzlich geregelt. Ähnlich wie Gebäude mit Rampen und Aufzügen ausgestattet sind, sollten auch Websites und Anwendungen durch alternative Eingabemethoden, Screenreader-Kompatibilität und flexible Darstellungsformen allen Nutzenden offenstehen (Morville, 2004; ISO 9241-210, 2020). Aus wissenschaftlicher Sicht ist Barrierefreiheit ein integraler Bestandteil der UX. Untersuchungen zeigen, dass mangelnde Zugänglichkeit zu erheblichen Gebrauchstauglichkeitsproblemen und damit zu geringer Nutzer*innenzufriedenheit führt, insbesondere, wenn ganze Nutzer*innengruppen vom Nutzungserlebnis ausgeschlossen werden (Petrie & Kheir, 2007).

Glaubwürdigkeit. Die Glaubwürdigkeit und Vertrauenswürdigkeit einer Website oder eines digitalen Produkts sind essenziell für eine positive UX. Nutzende müssen darauf vertrauen können, dass die präsentierten Informationen korrekt und verlässlich sind. Faktoren wie professionelles Design, klare Kommunikation, transparente Datenschutzrichtlinien und glaubwürdige Inhalte tragen dazu bei, die Glaubwürdigkeit eines digitalen Angebots zu stärken (Morville, 2004). Glaubwürdigkeit bedeutet somit, dass

Nutzende das Gefühl haben, dem Produkt oder Anbieter vertrauen zu können, eine wesentliche Voraussetzung für eine positive UX, denn ohne Vertrauen werden Nutzende das Angebot meiden, egal wie gut es funktional gestaltet ist (Corritore, Kracher & Wiedenbeck, 2003).

Wertstiftung. Ein digitales Produkt sollte nicht nur für die Nutzenden, sondern auch für das Unternehmen oder die Organisation dahinter einen Mehrwert bieten. Im kommerziellen Bereich bedeutet das eine Steigerung der Kundenzufriedenheit und eine Verbesserung der Geschäftsergebnisse. Im gemeinnützigen Bereich muss das Produkt die Mission der Organisation unterstützen. Diese Dimension verbindet die UX direkt mit den strategischen Zielen einer Organisation (Morville, 2004).

Das UX-Honeycomb-Modell wurde gewählt, da es als praxisorientiertes Framework funktionale, emotionale und strategische Dimensionen der Nutzer*innenwahrnehmung systematisch abbildet. Damit bietet es einen deutlich umfassenderen Analyserahmen als klassische Gebrauchstauglichkeitsansätze, die meist nur auf Effizienz oder Verständlichkeit fokussiert sind (Morville, 2004). Gerade für die Bewertung eines KI-Suchtools auf einer beratungsorientierten Reiseplattform wie Globetrotter ist diese ganzheitliche Perspektive entscheidend. Das UX-Honeycomb-Modell ist somit besonders geeignet, um nicht nur technische oder funktionale Stärken und Schwächen, sondern auch emotionale und kontextbezogene Nutzer*innenbedürfnisse sowie Akzeptanzhürden sichtbar zu machen. Dies ist essenziell, um gezielt Ansatzpunkte für die Verbesserung der UX eines KI-Suchtools zu identifizieren und die Forschungsfrage dieser Arbeit umfassend zu beantworten.

Methodisch wurde das Modell in mehreren Schritten operationalisiert. Es diente als Grundlage für die Entwicklung des Leitfadens der Nachbefragung und für die Bildung der Hypothesen im Nutzer*innentest. Ausserdem floss es in die Konzipierung des Kategoriensystems ein und bildete die Basis der qualitativen Inhaltsanalyse. Durch diese systematische Verankerung konnte sichergestellt werden, dass die Erhebung und Auswertung der Daten theoriegeleitet und praxisrelevant erfolgen. So konnten die Ergebnisse direkt in konkrete Handlungsempfehlungen überführt werden.

3 Methodik

Um die Forschungsfrage zu beantworten, wurde ein qualitatives, menschenzentriertes Forschungsdesign gewählt. Ziel ist es, herauszufinden, wie ein KI-Suchtool auf der Globetrotter-Website gestaltet sein muss, um die Bedürfnisse der Nutzer*innen zu erfüllen und eine positive UX zu ermöglichen. Die methodische Herangehensweise orientiert sich dabei eng an den Prinzipien der menschenzentrierten Gestaltung ISO 9241-210, an dem UX-Honeycomb-Modell und dem TAM. Diese Konzepte liefern die theoretische Grundlage für die Entwicklung der Erhebungsinstrumente und die anschliessende Analyse.

Neben der Durchführung der Nutzer*innentests mit der Think-Aloud-Methode wird in diesem Kapitel der Pretest sowie Sampling der Testpersonen erläutert. Zudem wird die Entwicklung des Interviewleitfadens, der Aufgaben der Nutzer*innentests sowie die angewandte qualitative Auswertungsmethode vorgestellt.

3.1 Nutzer*innentests

Die Bedürfnisse der Nutzenden wurden in sieben etwa 45-minütigen Remote-Nutzer*innentests erfasst. Ziel war es, Schwachstellen der bestehenden Lösung zu identifizieren, Herausforderungen aufzudecken und Optimierungspotenziale aufzuzeigen. Gemäss Moser (2012) reichen zwischen fünf und zehn Testpersonen für qualitative Untersuchungen aus.

Remote-Nutzer*innentests wurden gewählt, um sicherzustellen, dass sich die Testpersonen in ihrer gewohnten Umgebung befinden und ihre eigenen Geräte verwenden. Ausserdem konnten so sowohl der Bildschirm als auch die Kamera der Teilnehmenden aufgezeichnet werden. Dadurch wird die künstliche Situation von Nutzer*innentests reduziert und eine authentischere Nutzungssituation geschaffen. Dies trägt dazu bei, dass Testpersonen sich wohler fühlen, da sie persönliche Informationen eher zu Hause als in einer Laborumgebung preisgeben (De Bleecker & Okoroji, 2018).

Ausserdem entdecken Remote-Nutzer*innentests ungefähr gleich viele Gebrauchstauglichkeitsprobleme wie solche, die in einem Labor durchgeführt werden (Thompson, Rozanski & Haake, 2004). Gemäss Untersuchungen von Brush, Ames und Davis (2004) sowie Thompson, Rozanski und Haake (2004) bevorzugen die meisten Testpersonen einen Remote-Testansatz und fühlen sich dabei mindestens gleich wohl, ihre Gedanken laut zu äussern wie in einem Nutzer*innentest in einem Labor.

Die Nutzer*innentests wurden auf schweizerdeutsch durchgeführt, dies ermöglichte eine barrierefreie Kommunikation mit den Testpersonen. Alle Nutzer*innentests wurden aufgezeichnet, um die Beobachtungen zu dokumentieren und nachträglich zu analysieren.

Auf quantitative Methoden wurde bewusst verzichtet, da der Fokus dieser Arbeit auf subjektiven Meinungen, individuellen Erfahrungen und tiefgehenden Einblicken in die

Nutzer*innenperspektiven liegt. Qualitative Methoden bieten in diesem Kontext eine höhere Aussagekraft und ermöglichen ein umfassenderes Verständnis der subjektiven Meinungen und Erfahrungen der Testpersonen (Helfferich, 2011).

Besonderes Augenmerk wurde auf die Einhaltung ethischer Aspekte gelegt. Vor Testbeginn wurden die Teilnehmenden über den Zweck der Studie, die geplante Aufzeichnung, die Anonymisierung der Daten sowie den Ablauf des Tests instruiert (vgl. Anhang A & B).

3.1.1 Think-Aloud-Methode

Um die Gedanken der Testpersonen bestmöglich zu dokumentieren, wurden sie gebeten, die *Think-Aloud-Methode* anzuwenden. Während der Testdurchführung äusserten die Testpersonen ihre Gedanken laut. Dadurch konnten nicht nur konkrete Nutzungsprobleme aufgedeckt, sondern auch tiefere Einsichten gewonnen werden, die durch reine Beobachtung verborgen geblieben wären. Zeigten die Testpersonen während des Tests Phasen stillen Handelns, wurden sie höflich daran erinnert, ihre Gedanken erneut zu verbalisieren. Diese Methode birgt gemäss Nielsen (2012) den Nachteil, dass es für einige Personen ungewohnt oder unangenehm sein kann, während der Interaktion mit einem System durchgehend laut zu denken. Dennoch wurde diese Vorgehensweise gewählt, da sie leicht umzusetzen ist und mit geringem Aufwand einen hohen Erkenntnisgewinn ermöglicht (Nielsen, 2012).

3.1.2 Pretest

Um den entworfenen Nutzer*innentest zu verbessern und zu validieren, wurde ein Pretest mit einer Person aus dem Umfeld des Autors durchgeführt. Der Pretest ermöglichte es, Hypothesen, Aufgaben sowie die Fragen der Vor- und Nachbefragung gezielt zu optimieren. Zudem liess sich so die Länge des Nutzer*innentests realistisch einschätzen und die Moderation praxisnah erproben. Ursprünglich wurde überlegt, in der Einverständniserklärung auf den Begriff *KI* zu verzichten, um ein Priming zu vermeiden. Der Pretest machte jedoch deutlich, dass die Testpersonen bereits durch die Vorbefragung für das Thema sensibilisiert wurden, sodass der Verzicht auf den Begriff in der Einverständniserklärung keinen zusätzlichen Einfluss gehabt hätte.

3.1.3 Leitfadenerstellung Vor- und Nachbefragung

Für die Erstellung der Interviewleitfäden für die Vor- und Nachbefragung wurde zunächst auf Basis der Forschungsfrage sowie den vorgestellten Theorien die zentralen Themenblöcke gebildet. Für die Vorbefragung (vgl. Anhang C) standen dabei vor allem die Nutzer*innenbedürfnisse, mentalen Modelle und die Einstellungen gegenüber KI im Fokus. Ziel war es, ein fundiertes Verständnis der individuellen Erwartungen, Vorerfahrungen und

des Nutzungskontexts der Testpersonen zu gewinnen, Aspekte, die laut ISO 9241-210 (2020) essenziell sind.

Im nächsten Schritt wurden für die Nachbefragung (vgl. Anhang D) gezielt Themenblöcke entwickelt, die jeweils eine der sieben Dimensionen des UX-Honeycomb-Modells (vgl. Kap. 2.3.1) abbilden. Zu jeder Facette des UX-Honeycomb-Modells wurden jeweils spezifische Interviewfragen formuliert, um das KI-Suchtool zu evaluieren. Auf diese Weise konnte die UX des KI-Suchtools aus unterschiedlichen Perspektiven umfassend bewertet werden.

Die Ausarbeitung der Interviewfragen folgte dem SPSS-Prinzip (vgl. Abb. 4) nach Helfferich (2011). Alle Fragen wurden gesammelt, geprüft, sortiert und subsumiert. Durch diese methodische Herangehensweise wurde sichergestellt, dass sowohl die Bedürfnisse der Testpersonen als auch die UX-Bewertung theoriegeleitet und systematisch erhoben wurden. Dabei orientierte sich die Erhebung an der ISO 9241-210 für menschenzentrierte Gestaltung, dem TAM und den Dimensionen des UX-Honeycomb-Modells.



Abbildung 4. SPSS-Prinzip nach Helfferich (2011), eigene Darstellung. Das Prinzip visualisiert den Prozess der qualitativen Interviewleitfadenerstellung durch die vier Schritte Sammeln, Prüfen, Sortieren und Subsumieren

Innerhalb jedes Themenblocks wurden zunächst möglichst viele potenziell relevante Fragen gesammelt. Die genaue Formulierung oder inhaltliche Relevanz spielte in dieser ersten Phase noch keine Rolle. Im Anschluss erfolgten eine strukturierte Sichtung und Reduktion der Fragen. Redundante oder thematisch unpassende Fragen wurden entfernt, sodass letztlich nur diejenigen Fragen übrigblieben, die klar auf die Forschungsfrage sowie die Dimensionen der zugrunde liegenden Theoriemodelle abzielten.

Die mentalen Modelle und Voreinstellungen der Testpersonen im Hinblick auf KI wurden bewusst im Rahmen der Vorbefragung, also vor der Interaktion mit dem KI-Suchtool, erhoben. Ziel dieser Herangehensweise war es, ein möglichst authentisches und unbeeinflusstes Bild davon zu gewinnen, welche Vorstellungen, Erwartungen oder auch Vorbehalte die Testpersonen gegenüber KI-Tools mitbringen.

Dem Autor ist bewusst, dass diese Vorgehensweise die nachfolgenden Tests potenziell beeinflusst haben könnte, da die Aufmerksamkeit der Testpersonen gezielt auf bestimmte Themen, wie beispielsweise KI, gelenkt wurde. Dennoch wurde dieser methodische Kompromiss bewusst in Kauf genommen, da die unverfälschte Erhebung der mentalen Modelle als zentrale Grundlage für die Interpretation der Nutzungserfahrungen betrachtet wurde.

Nur durch diese Abfolge liessen sich vorher bestehende Vorstellungen klar von Eindrücken trennen, die erst im Verlauf der Nutzung entstanden sind, was im Hinblick auf die Forschungsfrage als entscheidend erachtet wurde.

3.1.4 Testaufgaben

Für die Konzipierung der Testaufgaben wurden auf Basis der Forschungsfrage, der ISO 9241-210, des UX-Honeycomb-Modells und des TAM, zentrale Bereiche festgelegt. Ergänzt wurden diese um die Dimensionen Interaktion, mentale Modelle, Nutzer*innenbedürfnisse und Herausforderungen, um alle für die Forschungsfrage relevanten Aspekte umfassend abzudecken. Aus diesen Bereichen wurden gezielt Hypothesen abgeleitet, die als Grundlage für die Entwicklung der Testaufgaben dienten. Die Übersicht aller Hypothesen sowie deren Bewertung ist im Anhang E dargestellt. Aufgrund ihres Umfangs wurde die Tabelle nicht im Fliesstext, sondern im Anhang platziert. Nach Durchführung der Nutzer*innentests wurden die Hypothesen im Ergebnisteil anhand der empirischen Erkenntnisse überprüft und entsprechend angenommen oder abgelehnt.

Bei der Entwicklung der Testaufgaben wurde darauf geachtet, die Authentizität der Nutzungssituation zu wahren und eine künstliche Lenkung der Testpersonen zu vermeiden. So wurde etwa Aufgabe 2 (vgl. Anhang F) bewusst offen und explorativ formuliert. Die Testpersonen sollten Angebote für ihr nächstes Reiseziel auf der Globetrotter-Website suchen, ohne expliziten Verweis auf das KI-Suchtool. Dadurch konnten die Testpersonen selbst entscheiden, welchen Navigationsweg sie wählen, sei es über die Kategorien oder das KI-Suchtool. Diese offene Aufgabenstellung ermöglichte es, das natürliche Nutzungsverhalten und insbesondere die Auffindbarkeit des KI-Suchtools realistisch zu beobachten.

Alle Aufgaben waren mit klaren Abschlusskriterien versehen, um die Zielerreichung objektiv beurteilen zu können. Ergänzend wurden Bonusaufgaben vorbereitet, um bei Bedarf weitere Aspekte der Interaktion zu erfassen (vgl. Anhang F).

Ziel der Konzipierung und Auswertung der Testaufgaben war es, Nutzer*innenbedürfnisse und potenzielle Herausforderungen systematisch zu identifizieren und daraus Handlungsempfehlungen zur Verbesserung der UX des KI-Suchtools auf der Globetrotter-Website abzuleiten.

3.2 Sampling

Das Sampling umfasste sieben Testpersonen, die als potenzielle Kund*innen von Globetrotter eingestuft werden können. Ziel war es, im Sinne der ISO 9241-210, Personen einzubeziehen, für die das Angebot von Globetrotter grundsätzlich relevant ist. Nur durch die Untersuchung realer, potenzieller Nutzenden lassen sich valide Rückschlüsse auf die tatsächliche UX des KI-Suchtools ziehen (Moser, 2012).

Um eine differenzierte Analyse der spezifischen Nutzer*innenbedürfnisse und Nutzungsherausforderungen zu ermöglichen, wurden gezielt zwei Merkmale in den Fokus gerückt. Erfahrungen mit Reiseagenturen sowie die KI-Affinität. Sie beeinflussen, welche Anforderungen und Erwartungen Nutzende an ein KI-Suchtool stellen, welche Herausforderungen bei der Nutzung auftreten und welche Handlungsempfehlungen daraus abgeleitet werden können. Diese Differenzierung stellte sicher, dass ein umfassendes Bild der Nutzer*innenbedürfnisse erhoben werden konnte. Auf eine explizite Gegenüberstellung der einzelnen Kategorien, Erfahrungen mit Reiseagenturen sowie KI-Affinität, wurde jedoch verzichtet. Zum einen waren die Unterschiede zwischen den Gruppen für die Beantwortung der Forschungsfrage nicht relevant, da das Ziel der Arbeit darin bestand, sämtliche Nutzer*innenbedürfnisse zu identifizieren und nicht um Unterschiede herauszuarbeiten. Zum anderen hätte eine detaillierte Analyse der Differenzen den Rahmen dieser Arbeit überschritten. Um sicherzustellen, dass die Testpersonen potenzielle Kund*innen von Globetrotter sind und um Unterschiede in ihrer KI-Affinität sowie ihrer Erfahrung mit Reiseagenturen erfassen zu können, füllten sie im Vorfeld der Nutzer*innentests einen Fragebogen aus (vgl. Anhang G).

Die Erfahrung mit Reiseagenturen wurde erfasst, um besser einschätzen zu können, welche Erwartungen die Testpersonen an ein KI-Suchtool haben könnten. Es ist denkbar, dass Nutzende mit mehr Agenturerfahrung gezieltere Ansprüche an Informationsstruktur und Beratung stellen. Weniger erfahrene Personen setzen hingegen möglicherweise andere Nutzungsschwerpunkte, etwa auf Eigenständigkeit oder Inspiration. Diese Differenzierung ermöglicht es, unterschiedliche Nutzungsprofile gezielt zu berücksichtigen und daraus passgenaue Handlungsempfehlungen für die Weiterentwicklung des KI-Suchtools abzuleiten.

Die KI-Affinität wiederum wurde erfasst, da Einstellungen gegenüber KI, technologische Selbstwirksamkeit sowie Vorerfahrungen mit vergleichbaren Technologien zentrale Einflussfaktoren für die Akzeptanz eines KI-Suchtools darstellen (Marangunic & Granić, 2015). Entsprechend beeinflussen diese Faktoren nicht nur die UX direkt, sondern geben auch Aufschluss darüber, in welchen Bereichen gezielt Transparenz, Unterstützung oder Gestaltungssicherheit geschaffen werden sollte. Personen mit wenig KI-Erfahrung

könnten sich von bestimmten Begrifflichkeiten oder Funktionsweisen eher überfordert fühlen, während erfahrene Nutzende möglicherweise höhere Erwartungen an Flexibilität, Personalisierung oder Antwortqualität stellen könnten.

Die gezielte Erhebung dieser beiden Merkmale ermöglicht eine differenzierte Analyse der UX und bildet die Basis, um spezifische Nutzer*innenbedürfnisse zu erkennen und darauf abgestimmte Handlungsempfehlungen abzuleiten. Die Testpersonen waren zwischen 23 und 60 Jahre alt, davon vier weiblich und drei männlich. Da Alter und Geschlecht für die Forschungsfrage keine zentralen Variablen darstellen, dienen sie lediglich der Kontextualisierung. Entscheidend war stattdessen die Relevanz des Nutzungsverhaltens für das KI-Suchtool und dessen Nutzung im konkreten Anwendungskontext. Die Rekrutierung der Testpersonen erfolgte mittels selektivem Sampling aus dem erweiterten privaten Umfeld des Autors. Dabei wurde bewusst darauf geachtet, keine Personen mit zu engem persönlichem Verhältnis einzubeziehen, um eine möglichst unbeeinflusste Testumgebung zu gewährleisten und sozial erwünschtes Antwortverhalten zu minimieren.

Tabelle 1 gibt eine Übersicht über das Sampling der Nutzer*innentests.

Tabelle 1
*Sampling Nutzer*innentests*

Interview Nr.	KI-Erfahrung	Erfahrung Reiseagenturen	Offenheit Reiseagenturen
B01	Fortgeschritten	Gelegentlich	Ja
B02	Keine	Regelmässig	Ja
B03	Mittel	Regelmässig	Ja
B04	Mittel	Gelegentlich	Ja
B05	KI-Experte	Gelegentlich	Ja
B06	Keine	Gelegentlich	Ja
B07	Fortgeschritten	Noch keine	Ja

3.3 Datenauswertung

Die Auswertung der im Rahmen der Nutzer*innentests erhobenen Daten erfolgte in einem mehrstufigen, theorie- und datenbasierten Verfahren. Die Videodateien der Tests, einschliesslich der Vor- und Nachbefragungen, wurden mithilfe der Software MAXQDA, angelehnt an Rädiker und Kuckartz (2019), auf Hochdeutsch transkribiert. Anschliessend wurde eine inhaltlich strukturierende Inhaltsanalyse nach Kuckartz und Rädiker (2020) durchgeführt.

Die Entwicklung des Kategoriensystems basierte neben der übergeordneten Forschungsfrage insbesondere auf dem Leitfaden für die Vor- und Nachbefragung sowie

auf den Hypothesen der Testaufgaben, die wiederum auf den Theorien der ISO 9241-210, des TAM und des UX-Honeycomb-Modells beruhen. Die zentralen Kategorien wurden somit deduktiv direkt aus den Dimensionen dieser Modelle, den Hypothesen der Tests sowie den Leitfadenfragen abgeleitet. Ziel war es, alle Aspekte von Nutzer*innenbedürfnissen, Nutzungserfahrungen und Herausforderungen in Bezug auf das KI-Suchtool umfassend zu erfassen.

Im Verlauf der Analyse wurde das deduktive Kategoriensystem um induktive Sub-Kategorien ergänzt, die empirisch deutlich hervortraten und von den theoretischen Kategorien nicht ausreichend erfasst wurden. Dazu gehörten vor allem spezifische Nutzer*innenbedürfnisse, Aspekte der konkreten Interaktion mit dem KI-Suchtool, Herausforderungen im Anwendungskontext sowie identifizierte Optimierungspotenziale.

Darüber hinaus wurde die Kategorie *Verschiedenes* eingeführt, in der zunächst Textpassagen gesammelt wurden, die beim ersten Codierdurchlauf in keine bestehende Kategorie passten. Diese Passagen wurden im weiteren Verlauf entweder einer passenden Kategorie zugeordnet oder, falls als nicht relevant eingestuft, aus der Analyse ausgeschlossen. Eine weitere induktive Kategorie bildete *Blumen am Wegesrande* als Sammelbecken für besonders spezifische, unerwartete oder zitierwürdige Aussagen.

Die qualitative Auswertung der Daten erfolgte in einem zweistufigen Codierprozess. In der ersten Phase wurden die transkribierten Textpassagen den deduktiv vom Leitfaden abgeleiteten Kategorien zugeordnet. Parallel wurden induktiv neue Sub-Kategorien gebildet, wann immer dies das Datenmaterial erforderte. Im zweiten Codierdurchlauf wurden sämtliche Textpassagen erneut geprüft und systematisch den induktiv entwickelten Kategorien zugeordnet. Zudem wurde sichergestellt, dass zuvor möglicherweise übersehene Passagen den relevanten deduktiven Kategorien nachträglich zugewiesen wurden. Diese iterative Vorgehensweise gewährleistete eine differenzierte, theoriegeleitete und zugleich empirisch fundierte Analyse der Nutzungserfahrungen mit dem KI-Suchtool.

Der Fokus der Analyse lag darauf, zentrale Herausforderungen zu identifizieren und daraus konkrete Handlungsempfehlungen abzuleiten. Die erkannten Herausforderungen und Optimierungspotenziale wurden systematisch ausgewertet und nach ihrer Problemschwere, der subjektiven Relevanz aus Sicht der Testpersonen und des Autors sowie der Häufigkeit der Nennungen gewichtet. Insgesamt wurden in der Auswertung 650 Textstellen codiert. Jede Kategorie wurde detailliert beschrieben und mit einem Ankerbeispiel versehen, um das Verständnis für die Bedeutung und Anwendung der jeweiligen Kategorie zu erhöhen und die Nachvollziehbarkeit der Analyse zu verbessern (vgl. Anhang H).

Zusammenfassend ermöglichte die Kombination aus deduktiv-theoriegeleiteten und induktiv-datenbasierten Kategorien ein umfassendes Verständnis der spezifischen Nutzer*innenbedürfnisse, Nutzungserfahrungen, Herausforderungen sowie Optimierungspotenzialen im Zusammenhang mit dem KI-Suchtool auf der Globetrotter-Website.

4 Ergebnisse

In diesem Kapitel werden die zentralen Erkenntnisse aus den qualitativen Nutzer*innentests vorgestellt. Die Ergebnisse liefern wertvolle Einblicke in die tatsächlichen Bedürfnisse und Erwartungen der Testpersonen im Umgang mit dem KI-Suchtool auf der Globetrotter-Website. Darüber hinaus wird der bestehende Prototyp systematisch hinsichtlich der sieben Dimensionen des UX-Honeycomb-Modells (vgl. Kap. 2.3.1) evaluiert. Abschliessend werden die aus den Nutzer*innentests gewonnenen Herausforderungen sowie Optimierungspotenziale aufgezeigt. Jede Aussage wird mit der entsprechenden Testperson als Quelle ausgewiesen, um sowohl die Nachvollziehbarkeit individueller Beiträge als auch die Frequenz bestimmter Äusserungen transparent zu machen.

4.1 Erfahrungen und Einstellungen

Dieses Kapitel gibt einen Überblick über die bisherigen Erfahrungen der Testpersonen mit KI sowie ihre grundlegende Haltung gegenüber KI-Tools. Zunächst werden die allgemeinen Erfahrungen und Einstellungen der Testpersonen zu KI-Tools dargestellt. Anschliessend werden zentrale Herausforderungen beschrieben, auf die Testpersonen bei der Verwendung klassischer Suchfunktionen auf Websites gestossen sind.

4.1.1 Erfahrungen und Einstellungen zu KI-Tools

Die allgemeine Einstellung der Testpersonen gegenüber KI-Tools reichte von sehr positiv bis skeptisch. Mehrere Testpersonen zeigten sich offen und begeistert und gaben an, KI-Tools bereits regelmässig einzusetzen, sei es im Arbeitsalltag oder privat, etwa zur Informationsrecherche oder Ideensammlung (B01; B03; B04; B05; B07). Zwei Testpersonen hatten KI-Tools bis anhin nicht bewusst verwendet (B02; B06) und standen ihnen eher skeptisch gegenüber. Unabhängig von der Nutzungsintensität betonten alle Testpersonen, man solle KI-Antworten kritisch hinterfragen und nicht blind übernehmen (B01 - B07). Insgesamt wurde KI zwar als hilfreich angesehen, etwa um schnell an Informationen zu gelangen, doch bestanden gleichzeitig Bedenken hinsichtlich Fehler und Missbrauchsmöglichkeiten von KI-Tools (B02; B03; B06; B07). Die grosse Mehrheit der Testpersonen sieht KI als unausweichlichen Trend der Zukunft, dem man sich vorsichtig annähern sollte (B01; B02; B03; B04; B05; B07).

4.1.2 Herausforderungen bei klassischen Suchfunktionen

Bei klassischen Suchfunktionen auf Websites stiessen alle schon auf Probleme. Häufig waren die gewünschten Informationen nicht auf Anhieb auffindbar oder die Suche ergab eine Vielzahl irrelevanter Treffer, wodurch die Identifizierung der relevanten Informationen erschwert wurde (B01 - B07). Mehrere Testpersonen bemängelten, dass

normale Suchfunktionen sehr wörtlich funktionieren. Gibt man nicht exakt den richtigen Begriff ein, erhält man unpassende Resultate (B01; B02; B03; B05; B07).

4.2 Mentale Modelle und Erwartungen

Dieses Kapitel beleuchtet zentrale mentale Modelle und Erwartungen der Testpersonen im Umgang mit Suchfunktionen auf Websites sowie mit dem KI-Suchtool auf der Globetrotter-Website. Zunächst werden die allgemeinen Erwartungen an klassische Suchfunktionen dargestellt, gefolgt von spezifischen Erwartungen an den KI-Kompass. Abschliessend wird aufgezeigt, welche typischen Suchstrategien Testpersonen beim Navigieren auf Websites verfolgen.

4.2.1 Erwartungen an eine Suchfunktion

Die Testpersonen hatten klare Erwartungen an die Suchfunktion einer Website. So erwarteten alle Testpersonen, schnell und unkompliziert gesuchte Informationen zu finden (B01 - B07). Gleichzeitig wurde betont, dass die Suchfunktion möglichst fehlertolerant sein sollte. Auch unvollständige oder unpräzise Eingaben sollten verstanden und sinnvoll interpretiert werden können (B01; B03; B05; B07). So äusserten Testpersonen die Erwartung, dass die Suchfunktion in der Lage sein sollte, auch zwischen den Zeilen zu lesen, also die Intention hinter einer Anfrage zu erkennen (B03; B05). Von zentraler Bedeutung war zudem, dass die Suchfunktion nicht nur relevante, sondern auch möglichst aktuelle Inhalte anzeigt (B01; B05). Bezüglich der Positionierung wurde von nahezu allen Testpersonen erwartet, dass eine Suchfunktion oben rechts auf der Website zu finden ist (B01; B02; B03; B04; B05; B07). Gleichzeitig hielten einige auch eine zentrale Platzierung auf der Startseite für sinnvoll, sofern diese auffällig und direkt nutzbar gestaltet ist (B01; B02; B05). Grundsätzlich wurde erwartet, dass eine Suchfunktion prominent platziert und intuitiv auffindbar sein sollte (B04; B06; B07).

Die Hypothese **H1**: *Nutzende erwarten eine Suchfunktion oben rechts auf einer Website zu finden* wird auf Grundlage der vorliegenden Ergebnisse angenommen.

4.2.2 Erwartungen an den KI-Kompass

Mehrere Testpersonen hatten die klare Erwartung, dass ihnen der KI-Kompass nicht nur allgemeine Informationen liefert, sondern konkrete Produktempfehlungen sowie direkte Buchungsoptionen anbietet (B01; B04; B05; B06; B07). Erwartet wurde auch, dass der KI-Kompass die Testpersonen aktiv durch den Suchprozess führt, gezielt Rückfragen stellt, um persönliche Reisebedürfnisse besser zu erfassen und individuellere Empfehlungen auszusprechen (B01; B04; B05; B07). Überraschung zeigte sich bei mehreren Testpersonen darüber, dass der KI-Kompass auch Fragen beantworten kann, die keinen direkten Reisebezug haben, etwa zu gesundheitlichen Themen oder allgemeinen Wissensfragen (B01; B03; B06; B07). Testpersonen, welche vor dem Nutzer*innentest

schon KI-Tools benutzt haben, stellten oft Vergleiche mit anderen KI-Tools auf (B01; B03; B04; B05; B07). Es wurde unter anderem erwartet, dass bereits geführte Chatverläufe auch bei einer neuen Suchanfrage gespeichert und abrufbar bleiben. Darüber hinaus wurde die Nutzung des KI-Kompass wiederholt mit der von ChatGPT verglichen (B01; B03; B04; B05).

4.2.3 Suchstrategien

Die Mehrheit der Testpersonen berichtete, bei der Suche zunächst die bestehenden Kategorien und Menüs zu nutzen und erst anschliessend, sofern sie dort nicht die gewünschten Informationen fanden, auf die Suchfunktion zurückzugreifen (B01; B02; B03; B04; B05; B07). Je nach Anliegen verwenden sie auch direkt die Suchfunktion (B01; B03; B04; B06). Wird die Suchfunktion verwendet, passen die Testpersonen ihre Suchbegriffe direkt an, wenn die angezeigten Ergebnisse nicht den Erwartungen entsprechen (B01 - B07). Dabei bevorzugen die Testpersonen in der Regel die Eingabe einzelner Schlagwörter oder Begriffe anstelle ganzer Sätze, um schneller ans Ziel zu gelangen (B01; B02; B04; B05). Auch hinsichtlich der verwendeten Geräte zeigte sich ein einheitliches Bild. Die Suche erfolgt hauptsächlich über das Smartphone oder den Laptop beziehungsweise PC (B01 - B07).

4.3 Nutzer*innenbedürfnisse

Dieses Kapitel beschreibt zentrale Bedürfnisse der Testpersonen im Umgang mit dem KI-Suchtool auf der Globetrotter-Website. Zunächst wird das Bedürfnis nach effizienter Informationsfindung beschrieben. Darauf folgt eine Betrachtung der kontextsensitiven und interaktiven Kommunikation sowie der Wunsch nach einer intuitiven und übersichtlichen Gestaltung. Abschliessend wird das Bedürfnis nach glaubwürdigen und aktuellen Ergebnissen thematisiert.

4.3.1 Effiziente Informationsfindung

Das zentrale Bedürfnis der Testpersonen bei der Verwendung einer Suchfunktion ist die schnelle und effiziente Auffindung relevanter Informationen. Alle Testpersonen äusserten explizit, dass sie es leid seien, sich durch irrelevante Treffer oder komplexe Navigationsstrukturen zu arbeiten und erwarten stattdessen, dass ihnen das KI-Suchtool gezielt und zeitsparend weiterhilft (B01 - B07). Besonders wichtig war der Wunsch, mit minimalem Aufwand und möglichst einer einzigen, klar formulierten Anfrage direkt zur gewünschten Information zu gelangen, ohne zahlreiche Anpassungen oder Umwege über Menüs (B01 - B07).

Eine Testperson formulierte ihr Bedürfnis prägnant mit den Worten: «So schnell wie möglich an mein Ziel zu kommen und das zu bekommen, was ich suche» (B02:19). Auch andere Testpersonen stimmten dem zu und betonten, dass ein KI-Suchtool ihnen idealerweise manuelle Zwischenschritte, wie etwa das Durchsuchen von Kategorien,

abnehmen sollte, um die Informationsbeschaffung zu vereinfachen (B02; B06; B07). Insgesamt zeigt sich, dass Effizienz und Schnelligkeit für die Testpersonen von herausragender Bedeutung sind (B01 - B07).

Die Hypothese **H2: Nutzende haben das Bedürfnis, mit dem KI-Suchtool schnell und gezielt personalisierte Reiseinformationen zu erhalten, ohne mehrere Suchanfragen tätigen zu müssen** wird auf Grundlage der vorliegenden Ergebnisse angenommen.

4.3.2 Kontextsensitive und Interaktive Kommunikation

Ein wichtiges Bedürfnis der Testpersonen bei der Verwendung des KI-Suchtools ist das Bedürfnis nach einer adaptiven und interaktiven Unterstützung im Suchprozess. Die Testpersonen erwarten, dass KI-Suchtools auch unvollständig, vage oder unpräzise formulierte Eingaben korrekt interpretieren können, ohne dass eine perfekte Formulierung erforderlich ist (B01 - B07). Dabei geht es nicht nur um die sprachliche Verarbeitung, sondern auch um die Fähigkeit, den Kontext und die dahinterliegenden Nutzer*innenbedürfnisse zu erkennen. Suchanfragen sollen nicht nur rein wörtlich verstanden, sondern inhaltlich interpretiert werden, sodass das System auch implizite Anliegen erfassen kann (B01; B03; B05; B07).

Darüber hinaus besteht ein starkes Bedürfnis nach einer interaktiven Führung durch den Suchprozess. Die Testpersonen wünschen sich, dass ein KI-Suchtool Rückfragen stellt oder passende Vorschläge macht, um die Suchintention besser zu erfassen und zu präzisieren (B01; B05; B07). Besonders hilfreich wird empfunden, wenn das KI-Suchtool in der Lage ist, frühere Eingaben zu berücksichtigen und darauf aufbauend kontextbezogene Anschlussinformationen oder personalisierte Empfehlungen zu liefern (B03; B04).

4.3.3 Intuitive und übersichtliche Gestaltung

Ein grundlegendes Anliegen der Testpersonen war die übersichtliche Struktur und Navigation der Website. Alle Testpersonen betonten die Relevanz einer klaren, logisch aufgebauten Seitenarchitektur, die eine intuitive Orientierung ermöglicht (B01 - B07). Dabei wurde besonders die Suchfunktion als wesentliches Element hervorgehoben, um gezielt und ohne Umwege an die gewünschten Informationen zu gelangen, insbesondere dann, wenn die Navigation über Kategorien als umständlich oder wenig zielführend empfunden wurde (B01; B03; B04; B05; B07). Erwartet wurde, dass die Suchfunktion gut sichtbar und leicht zugänglich ist, bevorzugt in der rechten oberen Ecke der Startseite (B01; B02; B03; B04; B05; B07).

Im Hinblick auf die Darstellung der Inhalte wurde der Wunsch nach klar verständlichen, kompakt aufbereiteten Informationen deutlich. Lange Textblöcke oder eine zu grosse Informationsdichte werden häufig als überfordernd empfunden. Stattdessen bevorzugen die

Testpersonen eine reduzierte, zielgerichtete Informationspräsentation, unterstützt durch eine angemessene Schriftgrösse, gut lesbare Schriftarten und eine klare visuelle Struktur (B01; B02; B03; B04; B05; B07).

Auch die allgemeine Gebrauchstauglichkeit war ein wiederkehrendes Thema. Die Testpersonen legten Wert auf eine einfache, nachvollziehbare Nutzung der Seite. Dazu zählt auch die Möglichkeit, jederzeit unkompliziert zur vorherigen Ansicht oder zur Ausgangslage, etwa einer gestarteten Suche, zurückzukehren. Deshalb sollte gemäss den Testpersonen, zentrale Funktionen und Orientierungspunkte so platziert sein, dass sie ohne langes Suchen zugänglich sind (B01; B04; B06; B07).

4.3.4 Glaubwürdige und aktuelle Ergebnisse

Viele Testpersonen äusserten das klare Bedürfnis nach verlässlichen, aktuellen und relevanten Informationen, wenn sie digitale Suchfunktionen nutzen. Insbesondere bei der Reiseplanung besteht der Wunsch, sich auf die angezeigten Inhalte vollständig verlassen zu können, sei es bei Angaben zu Einreisebestimmungen, empfohlenem Equipment oder saisonalen Besonderheiten. Die Ergebnisse einer Suchfunktion sollten deshalb nicht nur schnell auffindbar, sondern vor allem inhaltlich korrekt und auf dem neuesten Stand sein (B01; B02; B03; B04; B05; B07).

Eng damit verknüpft ist der Wunsch nach Transparenz hinsichtlich der Herkunft der Informationen. Viele Testpersonen möchten nachvollziehen können, auf welchen Quellen die Antworten beruhen und wer für deren Richtigkeit einsteht (B01; B04; B06; B07). Ebenso wurde von allen Testpersonen der offene Umgang mit den Grenzen digitaler Systeme positiv aufgenommen, etwa dann, wenn deutlich gemacht wird, dass Informationen nicht immer vollständig oder fehlerfrei sein können (B01 - B07).

Zusammenfassend lässt sich sagen, dass die Testpersonen ein starkes Bedürfnis nach Verlässlichkeit, Transparenz und Aktualität bei digitalen Suchfunktionen haben (B01 - B07).

4.4 Interaktion mit dem KI-Kompass

Im folgenden Kapitel werden zunächst die ersten Reaktionen der Testpersonen auf den KI-Kompass beschrieben. Anschliessend folgt eine Analyse des Suchverhaltens mit dem KI-Kompass, bevor abschliessend der Umgang mit Links thematisiert wird.

4.4.1 Erste Reaktion auf den KI-Kompass

Beim ersten Kontakt mit dem KI-Kompass fielen den Testpersonen sofort bestimmte Aspekte ins Auge. Die Erstreaktionen waren überwiegend positiv. Viele empfanden das Erscheinungsbild des Tools als modern, schlicht und ansprechend (B01; B02; B04; B05; B06; B07). Die zentrale Platzierung des KI-Kompasses in der Mitte der Startseite wurde mehrheitlich als positiv wahrgenommen, sie wurde von mehreren Testpersonen als direkt sichtbar und auffällig beschrieben (B01; B03; B05; B07). Allerdings übersahen einige das

KI-Suchtool zunächst oder waren unsicher, wie sie es benutzen sollten (B02; B06). Einige Testpersonen begannen ihre Suche direkt mit dem KI-Suchtool (B01; B02; B05; B07), während andere zunächst die Kategorien nutzten und erst anschliessend zum KI-Suchtool wechselten (B03; B04; B06).

Insgesamt war die anfängliche Haltung gegenüber dem KI-Kompass durch Interesse geprägt, jedoch auch von Unsicherheiten begleitet, ob und wie hilfreich die Antworten tatsächlich sind (B03; B05; B06; B07). Nach Erhalt der ersten Antworten waren viele positiv überrascht von der Geschwindigkeit sowie Qualität der Informationen (B04; B06; B07). Gleichzeitig wurden jedoch auch erste Kritikpunkte geäußert, insbesondere zu knapp gehaltenen oder unpräzisen Antworten (B01; B04; B05; B07).

4.4.2 Suchverhalten mit dem KI-Kompass

Im Verlauf des Tests zeigten die Testpersonen charakteristische Interaktionsmuster mit dem KI-Kompass. Die Mehrheit der Testpersonen benutzte zuerst nur Schlagwörter (B02; B03; B04; B05; B06), während andere direkt einen vollständigen Satz formulierten (B01; B07). Testpersonen, die zunächst mit Schlagwörtern suchten, formulierten ihre Anfragen später als ganze Sätze (B02; B03; B04; B05; B06). Bei einer ersten unzureichenden Antwort wurden direkt Folgefragen eingegeben (B01 - B07). Testpersonen mit KI-Vorkenntnissen zeigten sich insgesamt experimentierfreudiger (B01, B05). So stellte B05 aus persönlichem Interesse nach dem Nutzer*innentest noch weitere Fragen an das KI-Suchtool. Weniger erfahrene Testpersonen blieben eher bei klassischen, reiserelevanten Anfragen (B02, B06). Gemeinsam war allen Testpersonen, dass sie versuchten, ihr Ziel über ein dialogisches Frage-Antwort-Verfahren zu erreichen (B01 - B07). Testpersonen stellten eine erste Frage, bewerteten für sich die Antwort und reagierten anschliessend mit einer präzisierenden Anschlussfrage (B01; B03; B04; B05; B07).

Die Hypothese **H3**: *Nutzende verwenden überwiegend einzelne Schlüsselwörter oder kurze Begriffsfolgen bei der Verwendung des KI-Suchtools* wird auf Grundlage der vorliegenden Ergebnisse abgelehnt.

4.4.3 Umgang mit Links

Die in den Antworten enthaltenen Links wurden von allen Testpersonen aktiv genutzt (B01 - B07). Testpersonen hoben hervor, dass es ein deutlicher Mehrwert sei, wenn der KI-Kompass nicht nur eine komprimierte Antwort liefert, sondern gezielt auf tieferegehende Informationen verweist (B01; B02; B03; B05, B06). B01 und B03 lobten ausdrücklich, über die eingebetteten Links auch Erfahrungsberichte oder konkrete Reiseangebote entdecken zu können. Auch wenn sich der Fokus dabei vorübergehend vom Chatfenster weg

verlagerte, kehrten die Testpersonen danach wieder zurück, um neue oder weiterführende Fragen zu stellen.

Den in Abbildung 5 markierten Link zu den Datenschutz- und Nutzungshinweisen nahmen die meisten Testpersonen zwar wahr, hätten ihn bei privater Nutzung aber nicht angeklickt (B01; B02; B03; B05; B06). B05 gab offen an, keine Lust gehabt zu haben, AGBs durchzulesen und vermutete lediglich lange Rechtstexte hinter dem Link. B04 sowie B07 bemerkten den Hinweis erst, nachdem die Testleitung darauf hingewiesen hatte.



Abbildung 5. Der Datenschutz-Hyperlink wurde von den Testpersonen nicht selbstständig genutzt. Im Nachhinein bewerteten sie den Link jedoch als äusserst positiv (B01 - B07). Verfügbar unter: <https://www.globetrotter.ch/>

Da die Testleitung im Verlauf des Tests jedoch explizit darum bat, öffneten alle Testpersonen die entsprechende Seite. Nach dem Besuch äusserten sich die Testpersonen positiv überrascht. Die Informationen wurden als klar, hilfreich und aufschlussreich bewertet (B01 - B07). B05 korrigierte die erste Einschätzung und bewertete den Link nach dem Durchlesen als nützlich. Obwohl der Datenschutzhinweis im spontanen Nutzungskontext keine zentrale Rolle spielte, wurde seine Präsenz rückblickend von mehreren Testpersonen als sinnvoll und vertrauensfördernd eingestuft (B02; B03; B05; B07).

4.5 Evaluation des KI-Kompasses anhand des UX-Honeycomb-Modells

Zur systematischen Analyse der Nutzer*innenwahrnehmung wurde der KI-Kompass entlang der sieben Dimensionen des UX-Honeycomb-Modells (vgl. Kap. 2.3.1) ausgewertet. Die nachfolgenden Kapitel fassen die Ergebnisse dieser Dimensionen zusammen und

zeigen, wie der KI-Kompass aus Sicht der Testpersonen wahrgenommen wurde. Dabei stehen die tatsächlichen Erfahrungen und Aussagen der Testpersonen im Mittelpunkt, ohne Interpretation oder Bewertung. Ziel ist es, ein detailliertes Bild davon zu erhalten, welche Qualitäten das KI-Suchtool bereits erfüllt und in welchen Bereichen aus Sicht der Testpersonen Weiterentwicklungspotenzial besteht.

4.5.1 Nützlichkeit

Alle Testpersonen empfanden den KI-Kompass als hilfreiches Werkzeug, insbesondere um schneller an Informationen zu gelangen, als es mit einer klassischen Suchfunktion möglich wäre (B01 - B07). Viele zeigten sich positiv überrascht davon, wie effizient sie Antworten auf ihre Fragen erhielten. So stellte B03 fest: «Ich habe das Gefühl, du kommst definitiv viel schneller an die Sachen, weil du nicht so genau eingeben musst, was du suchst» (B03:119). Die Fähigkeit des Suchtools, vollständige Fragen zu verstehen und im besten Fall den Kontext zu erfassen, wurde von mehreren Testpersonen als grosser Vorteil gewertet (B01; B03; B04; B05; B07). Einige Testpersonen betonten, dass durch die dialogartige Interaktion präzisere Anfragen möglich seien, als es bei einer klassischen Suchfunktion der Fall wäre (B01; B02; B04). Nur eine Testperson zeigte sich zurückhaltender und meinte, in der aktuellen Version sei der Nutzen gegenüber der klassischen Suchfunktion noch nicht signifikant höher (B05). B05 begründete diese Einschätzung wie folgt:

Weil es mir zu wenig spezifische Antworten gibt und nicht wirklich das herauskommt, was ich suche. Ich meine klar, ich muss immer sagen, dass keine guten Antworten herauskommen, liegt an dem, der die Frage geschrieben hat, oder? Das ist ja klar, aber ich würde jetzt auch nicht irgendwie einen 20-zeiligen Prompt schreiben. Also es muss möglichst simpel und einfach sein und darum muss eigentlich auch ein schlechter Prompt zu einem guten Ergebnis führen können. (B05:98)

Allerdings merkte B05 auch an, dass dies eher am technologischen Reifegrad liege und sah für das Konzept grosses Potenzial in der Weiterentwicklung.

Zusammenfassend ergibt sich jedoch ein klares Bild. Aus Sicht der Testpersonen stiftet der KI-Kompass echten Mehrwert, da er Informationen schneller, bequemer und häufig auch in grösserem Umfang bereitstellt, als es über klassische Suchfunktionen möglich wäre (B01; B02; B03; B04; B06; B07).

Die Hypothese **H4**: *Nutzende empfinden das KI-Suchtool nützlicher als eine klassische Suchfunktion* wird auf Grundlage der vorliegenden Ergebnisse angenommen.

4.5.2 Benutzbarkeit

In Bezug auf die Benutzbarkeit wurde der KI-Kompass von den Testpersonen durchweg positiv bewertet. Die Anwendung wurde als intuitiv und unkompliziert empfunden (B01 - B07). Hervorgehoben wurde die einfache Eingabe in natürlicher Sprache ohne spezielle Suchoperatoren (B01; B03; B04; B05; B07).

Die Gestaltung der Benutzungsoberfläche wurde als klar und übersichtlich beschrieben. Es gibt keine überladenen Menüs, sondern lediglich zentrale Funktionen wie Frage eingeben, absenden und neue Suche starten (B01; B02; B04; B06). Auch wurde positiv angemerkt, dass der KI-Kompass tolerant gegenüber unpräziser oder umgangssprachlicher Eingabe ist. So wurde etwa gelobt, dass das KI-Suchtool auch dann passende Antworten liefert, wenn Fragen nicht exakt oder grammatikalisch korrekt formuliert sind (B01 - B07). Diese Flexibilität in der Interpretation der Eingaben wurde von mehreren Testpersonen als grosse Erleichterung beschrieben, da sie sich dadurch weniger Gedanken über die richtige Formulierung machen mussten (B02; B03; B07).

Insgesamt wurde die Benutzbarkeit des KI-Kompasses als sehr hoch eingeschätzt. Das KI-Suchtool ist niedrigschwellig, weitgehend selbsterklärend und funktioniert so, wie man es intuitiv erwartet (B01 - B07).

Die Hypothese **H5: Nutzende erkennen die Funktionsweise des KI-Suchtools intuitiv** wird auf Grundlage der vorliegenden Ergebnisse angenommen.

4.5.3 Attraktivität

Die visuelle Gestaltung des KI-Kompasses fand bei den Testpersonen grosse Zustimmung. Mehrere Testpersonen beschrieben das Design als modern, ansprechend und übersichtlich. Besonders positiv wurden dabei das Layout und das stimmige Farbschema bewertet, das sich an das Corporate Design von Globetrotter anlehnt. Besonders die dezente blaue Farbgebung wurde von mehreren Testpersonen als angenehm und professionell wahrgenommen (B01; B02; B03; B04; B06; B07). Die klare, nicht überladene Oberfläche erleichtere die Orientierung und lenke den Fokus auf die Inhalte, was laut den Rückmeldungen zur positiven Nutzer*innenwahrnehmung beitrug (B03; B07).

Auch der Name *KI-Kompass* wurde überwiegend als passend und einprägsam beurteilt, da er thematisch direkt mit Reisen assoziiert wird (B01; B02; B04; B06). Nur B07 gab an, sich unter dem Namen nicht wirklich etwas vorstellen zu können. Einzelne Kritikpunkte betrafen Details im UI-Design. So war für B04 die in Abbildung 6 dargestellte Suchleiste zu weit unten auf der KI-Kompass-Startseite.

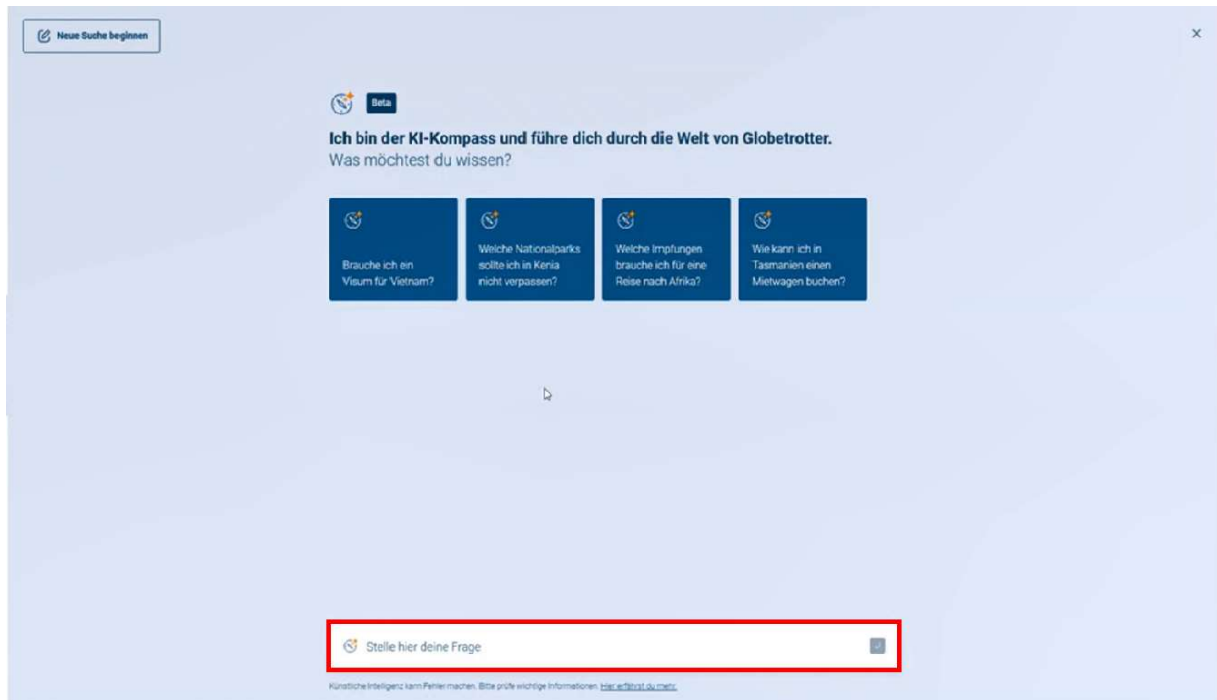


Abbildung 6. B04 merkte an, dass die rot markierte Suchleiste des KI-Kompasses zu weit unten ist und wünschte sich eine prominenterere Platzierung. Screenshot aus Nutzer*innentest B04

Eine Testperson empfand die in Abbildung 7 gezeigte Schriftgrösse auf der Informationsseite zu Datenschutz, Zweck und Limitationen als zu gross und visuell überfordernd (B01).



Abbildung 7. B01 empfand die Informationsseite über den KI-Kompass als visuell erdrückend. Verfügbar unter: <https://www.globetrotter.ch/ki-kompass-wie-funktioniert>

Als gelungen empfand eine Testperson hingegen subtile visuelle Hinweise auf die KI-Funktionalität, etwa durch die Verwendung von Sternsymbolen im Icon, die auf KI hinweisen (B05). Eine andere Testperson äusserte jedoch, dass das Kompass-Icon nicht sofort eindeutig sei, seine Bedeutung erschliesse sich erst im Zusammenspiel mit dem

Namen (B04). Bezüglich der Positionierung des Suchfelds äusserte sich eine Person, dass sie die zentrale Platzierung auf der Startseite als ästhetisch störend empfand (B07).

Insgesamt wurde die Attraktivität des KI-Kompasses als positiv bewertet. Der visuelle Auftritt wurde als ansprechend, funktional und gut auf den Nutzungskontext abgestimmt beschrieben, ohne auffällige Schwächen im Design (B01; B02; B04; B06).

Die Hypothese **H6: Nutzende mögen das schlichte Design des KI-Suchtools** wird auf Grundlage der vorliegenden Ergebnisse angenommen.

4.5.4 Auffindbarkeit

Die Auffindbarkeit wurde von den Testpersonen in zweifacher Hinsicht thematisiert. Zum einen ging es darum, wie leicht der KI-Kompass selbst auf der Website zu finden ist, zum anderen, wie gut sich mit seiner Hilfe relevante Inhalte auffinden lassen.

Bezüglich der Sichtbarkeit des Tools auf der Website waren die Erfahrungen gemischt. Einige Testpersonen nahmen den KI-Kompass direkt wahr, da er auf der Startseite gut sichtbar eingebunden war (B01; B04; B05; B07). Andere hingegen hätten ihn ohne Hinweis der Testleitung übersehen (B06). Insbesondere der Zugang über die Seitenlasche am rechten Bildschirmrand war nicht für alle intuitiv verständlich (B02; B04; B05).

Deutlich positiver fielen die Rückmeldungen zur inhaltlichen Auffindbarkeit aus. Mehrere Testpersonen berichteten, dass ihnen der KI-Kompass bereits mit der ersten Antwort eine strukturierte Übersicht zum Thema bot (B01; B07). Zentrale Aspekte wurden häufig genannt und ein gelungener Einstieg ins Thema geboten, der eine weitere Recherche ermöglichte (B03). Besonders hilfreich fanden die Testpersonen die eingebetteten Links, da sie gezielt zu weiterführenden Informationen führen und so eine schnellere Vertiefung ermöglichen als die Navigation über Kategorien (B01).

Einzelne Einschränkungen betrafen die Tiefe der direkt angezeigten Informationen. So wurde mehrfach angemerkt, dass die Antworten der KI eher oberflächlich seien und Details oft erst über zusätzliche Suchen erschlossen werden müssten (B01; B05). Eine Testperson beschrieb dies so:

Also, es gibt einem schnell eine Übersicht. Es ist aber nicht einfach gewesen, vertiefte Informationen zu finden, so ein wenig das Reasoning, das fehlt mir ein bisschen. Also die vertieften Informationen, die vielleicht nicht mal direkt ersichtlich im prompt sind, welche so ein bisschen darüber hinausgehen, also ein bisschen zu erkennen, was könnten die Bedürfnisse sein? Dort sehe ich noch Verbesserungspotenzial. (B01:81)

Trotz dieser Kritikpunkte wurde der KI-Kompass insgesamt als deutlich effektiver im Auffinden relevanter Inhalte wahrgenommen als eine klassische Suchfunktion. Viele empfanden es als grosse Erleichterung, dass sie nicht selbst erraten mussten, wo sich

bestimmte Informationen verbergen könnten, der KI-Kompass zeigte stattdessen klar erkennbare Pfade auf (B03; B04).

Die Hypothese **H7: Nutzenden fällt auf der rechten Seite der Website die KI-Kompass-Lasche nicht auf** wird auf Grundlage der vorliegenden Ergebnisse angenommen.

4.5.5 Zugänglichkeit

Die Zugänglichkeit des KI-Kompasses wurde von den Testpersonen grösstenteils positiv bewertet, insbesondere im Hinblick auf unterschiedliche technische Vorkenntnisse. Die Anwendung sei einfach gehalten und erfordere keine besonderen Vorkenntnisse oder technischen Fähigkeiten (B03; B06; B07). Auch Testpersonen mit geringer Technikaffinität könnten das KI-Suchtool problemlos nutzen, da die Bedienungsschritte klar verständlich seien und sich intuitiv erschliessen (B03; B06).

Als besonders nutzer*innenfreundlich wurde auch wahrgenommen, dass keine Registrierung oder Voreinstellungen notwendig sind. Das KI-Suchtool lässt sich sofort nutzen, was die Einstiegshürde deutlich senkt (B02).

Gleichzeitig wurde angemerkt, dass die Qualität der Ergebnisse davon abhängt, wie die Eingaben formuliert sind. Wer präzisere oder kontextreichere Fragen stellt, erhält relevantere Antworten. Diese Anforderung, sinnvolle und gezielte Fragen zu stellen, wurde als potenzielle Einstiegshürde für weniger geübte Nutzende benannt, allerdings nicht als grundsätzliches Zugänglichkeitsproblem, sondern als Teil einer Lernkurve, die sich durch Übung überwinden lässt (B01; B02; B03; B04; B05). Niemand hatte den Eindruck, dass bestimmte Nutzer*innengruppen ausgeschlossen werden. Im Gegenteil betonte eine Testperson, dass nichts passieren könne, selbst wenn man wenig Erfahrung habe, man müsse sich einfach trauen, das KI-Suchtool auszuprobieren (B04).

Auch die konkrete Nutzung, also eine Frage einzugeben und eine Antwort zu erhalten, wurde durchweg als unkompliziert und problemlos beschrieben.

Insgesamt wurde der KI-Kompass als niedrighürde und breit zugängliche Anwendung wahrgenommen, die sich ohne Vorkenntnisse oder technische Hürden bedienen lässt (B01 - B07).

4.5.6 Glaubwürdigkeit

Die Glaubwürdigkeit der vom KI-Kompass gelieferten Informationen wurde von den Testpersonen differenziert, insgesamt aber positiv eingeschätzt. Mehrere Testpersonen gaben an, den Antworten der KI grundsätzlich zu vertrauen, vor allem deshalb, weil die Inhalte erkennbar von Globetrotter stammen. Die Marke selbst wirkte dabei vertrauensstiftend (B01; B02; B03; B04; B05). Eine Testperson betonte, Globetrotter sei für

sie seit vielen Jahren ein seriöser Anbieter, was ihr Vertrauen in das KI-Suchtool stärkte (B02).

Die Testpersonen hoben hervor, wie wichtig ihnen die Transparenz der Informationsquellen ist. Besonders positiv wurde bewertet, wenn das System die zugrunde liegenden Daten offenlegt, zum Beispiel unternehmensinterne Erfahrungswerte, da dies das Vertrauen stärkt (B02; B03). Positiv fiel zudem auf, dass der KI-Kompass seine eigenen Grenzen kennt. Hinweise, dass KI auch Fehler machen kann oder der Verweis auf menschliche Beratung wurden als verantwortungsvoll bewertet (B01 - B07). Diese Offenheit wurde als Zeichen dafür verstanden, dass Globetrotter bemüht sei, die Qualität und Korrektheit der Auskünfte transparent zu machen (B03).

Trotz grundsätzlichem Vertrauen begegneten die Testpersonen den Antworten nicht unkritisch. Mehrere Testpersonen gaben an, sich bei Unsicherheiten zusätzlich absichern zu wollen, etwa durch einen Abgleich mit anderen Quellen (B02; B05; B07). In der Praxis war eine Überprüfung der Angaben jedoch nicht erforderlich, da die gelieferten Informationen überwiegend plausibel und mit bekanntem Wissen konsistent waren. Die korrekte Wiedergabe von Fakten trug zudem zur Stärkung von Vertrauen und Glaubwürdigkeit bei (B01; B06). Skeptisch reagierte eine Testperson auf offensichtlich erkennbare Eigenwerbung, etwa wenn das KI-Suchtool ausschliesslich Produkte von Globetrotter empfahl, wie im konkreten Fall eines bestimmten Reiserucksacks (B07). In solchen Situationen wurde die Objektivität der Empfehlung hinterfragt. Gleichzeitig wurde anerkannt, dass ein von Globetrotter betriebenes KI-Suchtool naturgemäss eigene Angebote in den Vordergrund stellt. Solange die Informationen dabei hilfreich und korrekt seien, wurde dies nicht als negativ bewertet (B07).

Datenschutzbedenken spielten kaum eine Rolle. Die meisten Testpersonen empfanden die Nutzung als unbedenklich, da keine persönlichen oder sensiblen Daten eingegeben werden mussten (B02; B04; B06; B07). Einige Testpersonen gaben an, keine privaten Informationen in ein KI-Tool einzugeben, fühlten sich durch die Nutzung des KI-Kompasses jedoch nicht in ihrer Datensouveränität eingeschränkt (B03). Vielmehr wurde die Anwendung im digitalen Alltag als unkritisch eingestuft (B01; B03; B04).

Insgesamt schätzten die Testpersonen den KI-Kompass als vertrauenswürdige Informationsquelle ein, unterstützt durch bekannte Markenautorität, transparente Kommunikation und realistische Einschätzungen der Systemgrenzen. Dieses Vertrauen wurde jedoch stets durch gesunden Menschenverstand begleitet und durch eine Bereitschaft zur Nachprüfung ergänzt (B01 - B07).

Die Hypothese **H9**: *Nutzende werden durch den Hinweis: «Künstliche Intelligenz kann Fehler machen. Bitte prüfe wichtige Informationen.» verunsichert* wird auf Grundlage der vorliegenden Ergebnisse abgelehnt.

4.5.7 Wertstiftung

Die Testpersonen identifizierten mehrere Funktionen des KI-Kompasses, die potenziell zur Wertstiftung für Globetrotter beitragen. Einige Testpersonen betonten, dass die vom KI-Suchtool gelieferten Antworten gezielt mit weiterführenden Links innerhalb der Globetrotter-Website versehen waren, etwa zu Reiseberichten oder Beratungsangeboten. Diese direkte Verknüpfung wurde als hilfreich und wertstiftend empfunden (B01; B02; B04; B06).

Einige Testpersonen waren überrascht, dass der KI-Kompass auch fachfremde oder Allgemeinwissen bezogene Fragen beantworten konnte. Dabei stellte das System stets einen Bezug zum Reisen her, zum Beispiel durch Hinweise auf gesundes Reisen oder saisonale Themen wie die Osterferien (B01; B02; B03; B07). So äusserte sich B02 wie folgt:

Das finde ich total witzig, dass es das gerade wieder mit Globetrotter, also mit dem Reisebüro in Verbindung bringt und dass da auch noch gerade eben so die Idee gibt, ja das eben für eine Reise zu nutzen. Das finde ich jetzt cool, dass es nicht einfach nur sagt, Ostern fällt auf den 20. April in diesem Jahr. (B02:66)

Die Zuordnung der Inhalte zur Marke Globetrotter wurde von mehreren Testpersonen bewusst wahrgenommen. Sie gaben an, den Informationen zu vertrauen, weil sie mit der Marke positive Erfahrungen verbinden oder Globetrotter als etablierten Anbieter wahrnehmen (B01; B02; B03; B04; B05).

Eine Testperson bemerkte, dass das KI-Suchtool gezielt eigene Produkte empfiehlt, wie zum Beispiel einen Reiserucksack. Diese Eigenwerbung wurde erkannt, teilweise kritisch gesehen, aber als zum Angebot gehörig akzeptiert (B07).

In einem Fall wurde angemerkt, dass der KI-Kompass bei einer konkreten Anfrage nach persönlicher Beratung lediglich allgemeine Empfehlungen oder sogar auf andere Reiseanbieter verwies und keine spezifischen Globetrotter-Dienstleistungen nannte (vgl. Abb. 8).



Abbildung 8. Der KI-Kompass gibt auf die Frage nach Beratung für Mallorca an, ein Reisebüro aufzusuchen und verweist auf Mallorca Tourismus. Screenshot aus Nutzer*innentest B05

Dadurch blieb eine mögliche Weiterleitung ins Beratungssystem aus, was die Testperson, bezüglich der Wertstiftung, als sehr kritisch empfand (B05).

Die Hypothese **H10: Nutzende empfinden die eingebetteten Links als wertstiftend** wird auf Grundlage der vorliegenden Ergebnisse angenommen.

4.6 Identifizierte Herausforderungen

In den Nutzer*innentests wurden verschiedene Herausforderungen deutlich, die die UX des KI-Kompasses einschränken. Mehrere Testpersonen gaben an, dass der Zugang zum KI-Kompass, insbesondere die am Seitenrand platzierte Öffnungslasche, nicht sofort ersichtlich war (B02; B05; B06). Auch Navigationsprobleme innerhalb des KI-Suchtools wurden beschrieben. Zwei Testpersonen berichteten, dass nach dem Start einer neuen Suche die vorherigen Dialogverläufe nicht mehr einsehbar waren. Dadurch war es nicht mehr möglich, auf bereits gestellte Fragen oder frühere Inhalte zurückzugreifen (B01; B04).

Eine weitere Herausforderung bestand darin, dass Testpersonen nach dem Anklicken eines Links mit der Zurück-Funktion (vgl. Abb. 9) statt zum Chat auf die Startseite zurückgeleitet wurden (B01; B02; B06).



Abbildung 9. Testpersonen haben erwartet, mithilfe des rot markierten Zurück-Buttons wieder auf den Chat mit dem KI-Kompass zu gelangen. Seite verfügbar unter: <https://www.globetrotter.ch/reise/mallorca-mit-dem-mietwagen-unterwegs>

Mehrere Testpersonen thematisierten inhaltliche Limitationen der KI-Antworten. Teilweise wurden die Auskünfte als zu kurz, oberflächlich oder nicht ausreichend spezifisch empfunden (B01; B03; B05; B07). Es fehlten beispielsweise weiterführende Informationen oder Hinweise, die für eine fundierte Entscheidung wichtig gewesen wären, etwa bei sicherheitsrelevanten Fragen zu Reisezielen (B02). In solchen Fällen mussten die Testpersonen Rückfragen stellen oder ihre Eingabe mehrfach überarbeiten, um die gewünschten Informationen zu erhalten (B01; B02; B07).

Technische Probleme wurden auch vereinzelt beobachtet. In einem Fall reagierte der KI-Kompass nicht auf die Eingaben der Testperson (vgl. Abb. 10). Auch das erneute Laden der Seite brachte keine Lösung. Die Testperson musste auf ein anderes Gerät wechseln, damit die Seite funktionierte (B04).

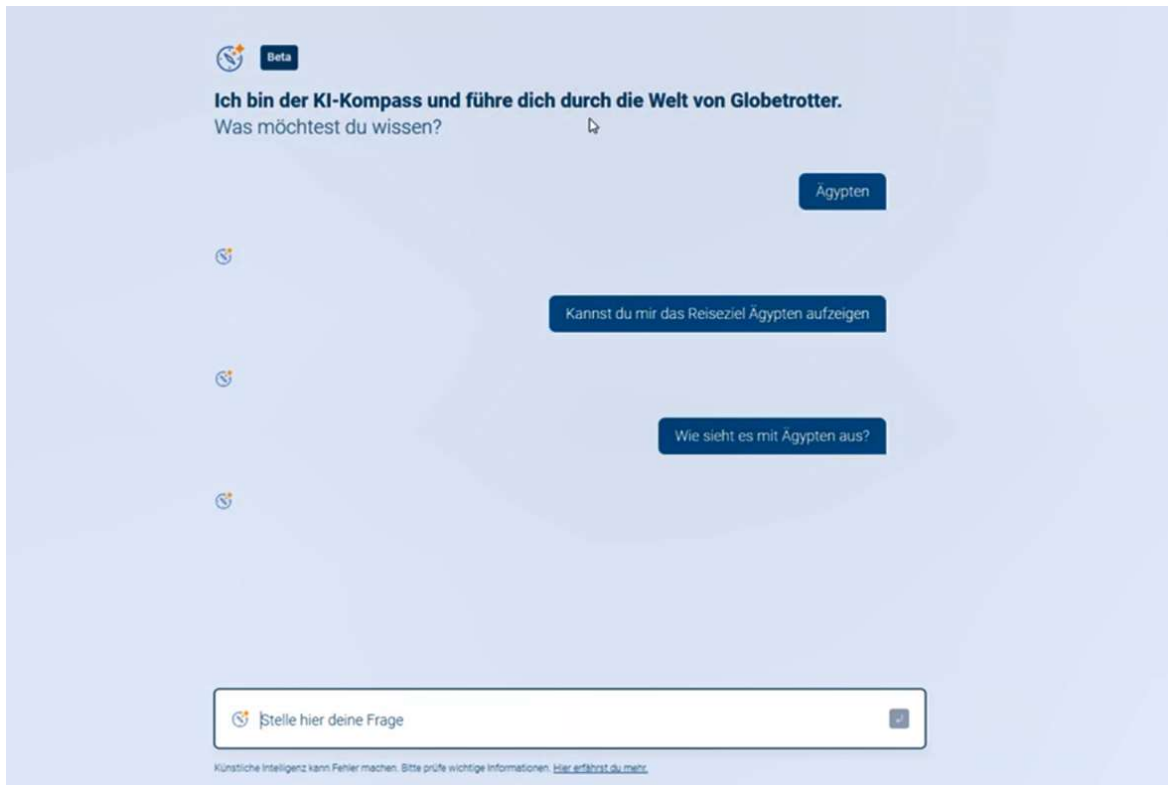


Abbildung 10. Der KI-Kompass reagiert nicht auf die Eingaben der Testperson. Screenshot aus dem Nutzer*innentest B04

In einem anderen Fall führte ein vorgeschlagener Link B02 zu einer Fehlermeldung, dass die Seite nicht gefunden wurde (vgl. Abb. 11).



Abbildung 11. Durch einen Link in einer Antwort bekam B02 eine Fehlermeldung. Screenshot aus Nutzer*innentest B02

Ausserdem stellte eine KI-erfahrene Testperson fest, dass die Generierung der Antworten viel Zeit in Anspruch nahm und vermutete als Ursache ein langsames Sprachmodell (B05).

Schliesslich wurden auch visuelle Ablenkungen auf der Website als störend wahrgenommen. Zwei Testpersonen berichteten, dass das weiterlaufende Video auf der Startseite beim Lesen der Antworten visuell ablenkte und ihre Konzentration beeinträchtigte (B03; B07).

Die Hypothese **H11**: *Nutzende fühlen sich frustriert, wenn das KI-Suchtool unpräzise oder nicht nachvollziehbare Ergebnisse liefert* wird auf Grundlage der vorliegenden Ergebnisse angenommen.

4.7 Identifizierte Optimierungspotenziale

Die Testpersonen äusserten konkrete Verbesserungsvorschläge zur Weiterentwicklung des KI-Suchtools. Ein häufig angesprochener Punkt betraf die Sichtbarkeit und Einbettung des Tools auf der Website. Mehrere Testpersonen gaben an, dass sie den KI-Kompass gerne prominenter platziert sehen würden, beispielsweise durch auffälligere Icons und Beschriftungen (B01; B05; B07). Eine Testperson schlug vor, dass klarer hervorgehoben werden sollte, dass es sich um ein KI-Suchtool handelt, damit Nutzende dies auf den ersten Blick erkennen (B01). Von einer weiteren Testperson wurde angeregt, direkt auf der Website darauf hinzuweisen, welche Arten von Fragen das KI-Suchtool beantworten kann (B07).

Eine Testperson hinterfragte, ob das ausklappbare Chatfenster als Form des KI-Kompasses wirklich notwendig ist. Sie regte an, das KI-Suchtool stattdessen direkt in die bestehende Suchleiste zu integrieren, vor allem, wenn es hauptsächlich zur Informationssuche und weniger für Beratung genutzt wird. Gleichzeitig sprach sie sich dafür aus, die Chatfunktion zu erhalten, falls das KI-Suchtool zukünftig für mehrstufige Interaktionen bis hin zur Buchung genutzt werden soll (B05).

Ein weiterer thematischer Schwerpunkt der Rückmeldungen betraf das Erwartungsmanagement und die Kommunikation der Funktionalität des KI-Suchtools. Mehrere Testpersonen gaben an, dass sie zu Beginn nicht genau wussten, was der KI-Kompass leisten kann oder was sie damit anfangen sollen (B02; B03; B06; B07). So teilte eine Testperson ihre Gedanken folgendermassen: «Was möchtest du von unserem KI-Kompass wissen? Das finde ich jetzt noch spannend. Ich weiss nicht genau, was damit gemeint ist» (B02: 31).

Eine Testperson berichtete, dass sich ihr der Name des KI-Suchtools erst durch die zusätzliche Anzeige *Stell mir eine Frage* erschlossen habe (B07). Von anderen wurde gewünscht, dass frühzeitig kommuniziert wird, welche Art von Antworten zu erwarten sind, etwa ob es sich um eine klassische Suchfunktion, eine digitale Beratung oder eine Mischform handelt (B04; B05; B07).

Hinsichtlich der Dialogführung und Antwortstruktur sprachen mehrere Testpersonen konkrete Vorstellungen aus. Sie äusserten den Wunsch, dass der KI-Kompass bei unklaren oder allgemein formulierten Fragen zunächst Rückfragen stellen sollte, bevor er eine Antwort gibt. Dadurch könnte das Anliegen genauer eingegrenzt werden (B01; B05; B07). Eine Testperson schlug vor, dass bei Eingabe eines Reiseziels direkt nach der Art der geplanten Reise gefragt werden sollte, zum Beispiel, ob es sich um Erholung, Abenteuer oder eine Rundreise handelt (B05). Auch die Idee, dass der KI-Kompass eigenständig passende Empfehlungen gibt, zum Beispiel konkrete Produktvorschläge mit einem Bild bei Fragen zu Reisezubehör, wurde angemerkt (B07).

Mehrere Testpersonen wünschten sich, dass frühere Eingaben und Dialoge gespeichert werden, da der bisherige Suchverlauf bei jeder neuen Suche gelöscht wird und somit nicht erneut aufgerufen werden kann (B01; B04). Eine Testperson führte aus, dass eine solche Verlaufsspeicherung insbesondere für längere Rechercheprozesse hilfreich wäre und bei anderen KI-Tools bereits Standard sei (B01).

Darüber hinaus nannten Testpersonen auch technische und inhaltliche Verbesserungsvorschläge. Eine Testperson regte an, die Antwortzeit der KI zu verkürzen, da lange Ladezeiten als störend empfunden wurden (B05). Auch der Wunsch nach ergänzenden Kontextinformationen wurde geäußert. Eine Testperson schlug vor, dass bei der Suche nach Reisezielen in politisch instabilen Regionen automatisch offizielle Reisehinweise, etwa vom Eidgenössischen Departement für auswärtige Angelegenheiten, angezeigt werden sollten (B02).

Insgesamt machten die Testpersonen deutlich, dass sie neben der Gebrauchstauglichkeit vor allem eine präzisere Informationsaufbereitung, gezielte Rückfragen und eine klarere Einbettung des Tools in die Website als mögliche Weiterentwicklungen ansehen.

Die Hypothese **H8**: *Nutzende wünschen sich eine kurze Erklärung zur Funktionsweise des KI-Suchtools* wird auf Grundlage der vorliegenden Ergebnisse angenommen.

5 Diskussion

Im folgenden Kapitel werden die empirischen Befunde im Kontext der vorgestellten Theorien diskutiert. Ziel ist es, daraus abzuleiten, wie das KI-Suchtool auf der Globetrotter-Website gestaltet sein muss, um die UX aus Sicht der Nutzenden nachhaltig zu verbessern. Ziel ist es, die wichtigsten Nutzer*innenbedürfnisse, die zentralen Herausforderungen des KI-Suchtools und die daraus abgeleiteten Handlungsempfehlungen übersichtlich zusammenzufassen.

5.1 Beantwortung der Fragestellung

Die Forschungsfrage dieser Arbeit lautet:

Welche spezifischen Nutzer*innenbedürfnisse bestehen im Zusammenhang mit einem KI-gestützten Suchtool auf der Globetrotter-Website, welche Herausforderungen ergeben sich bei dessen Nutzung und welche konkreten Handlungsempfehlungen lassen sich daraus ableiten, um die User Experience gezielt zu verbessern?

Zur systematischen Beantwortung dieser Frage werden die empirischen Ergebnisse der Nutzer*innentests mit den theoretischen Modellen und Grundlagen in Beziehung gesetzt. Die Forschungsfrage unterteilt sich in drei zentrale Aspekte, die in den folgenden Unterkapiteln ausführlich erläutert werden. Abschliessend erfolgt eine kompakte, zusammenfassende Beantwortung der Forschungsfrage. In den folgenden Kapiteln zu den Nutzer*innenbedürfnissen, den identifizierten Herausforderungen sowie den Handlungsempfehlungen orientiert sich die Reihenfolge der dargestellten Aspekte jeweils an ihrer Gewichtung. Diese Gewichtung basiert sowohl auf der quantitativen Häufigkeit der Nennungen im Rahmen der Nutzer*innentests (vgl. Anhang H) als auch auf der subjektiven Relevanz der einzelnen Punkte, wie sie von den Testpersonen und dem Autor wahrgenommen wurden. So wird sichergestellt, dass jeweils mit dem für die Nutzenden wichtigsten Bedürfnis, der zentralsten Herausforderung beziehungsweise der bedeutsamsten Handlungsempfehlung begonnen wird.

5.1.1 Nutzer*innenbedürfnisse

Effiziente Informationsfindung. Die durchgeführten Nutzer*innentests verdeutlichen, dass Nutzende bei KI-Suchtools vor allem Wert auf eine effiziente Informationsfindung legen (vgl. Kap. 4.3.1). Testpersonen äusserten den Wunsch, mit möglichst geringem Aufwand direkt zu den gesuchten Informationen zu gelangen. Besonders betont wurde dabei die Relevanz und Präzision der KI-Antworten (vgl. Kap. 4.3.1). Dieses Bedürfnis nach schneller, zielgerichteter Informationsbeschaffung steht im Einklang mit den Prinzipien menschenzentrierter Gestaltung, die Systeme so entwickeln will, dass sie effektiv und effizient nutzbar sind. Gemäss ISO 9241-210 (2020) sollen

Anwendungen die Aufgaben der Nutzenden optimal unterstützen, das bedeutet unter anderem, dass bereits eine kurze, direkte Sucheingabe zu brauchbaren Ergebnissen führen sollte. In Anlehnung an das TAM trägt eine solche wahrgenommene Nützlichkeit massgeblich zur Akzeptanz des Systems bei, da Nutzende die damit verbundene Effizienzsteigerung als vorteilhaft empfinden (Marangunić & Granić, 2015). Das Bedürfnis nach effizienter Informationsfindung ist gemäss dem UX-Honeycomb-Modell relevant, weil Effizienz direkt zur Dimension der Nützlichkeit beiträgt. Nur wenn ein KI-Suchtool Nutzende schnell und gezielt an die gewünschten Informationen bringt, wird es als wirklich nützlich erlebt und bietet somit einen echten Mehrwert für die UX (Morville, 2004).

Kontextsensitive und interaktive Kommunikation. Auch das Bedürfnis nach kontextsensitiver, interaktiver Kommunikation im Suchprozess wurde von vielen Testpersonen betont. Sie erwarten, dass ein KI-Suchtool auch unvollständige oder vage Anfragen inhaltlich interpretiert und bei Unklarheiten gezielt Rückfragen stellt, um die Nutzungsintention besser zu erfassen (vgl. Kap. 4.3.2). Dieses Bedürfnis steht im Einklang mit dem Prinzip der ISO 9241-210 (2020), den Nutzungskontext sorgfältig zu analysieren und die Nutzenden kontinuierlich in die Gestaltung einzubeziehen. Auch das UX-Honeycomb-Modell unterstreicht in der Facette Nützlichkeit, dass ein digitales Produkt einen klaren, erlebbaren Nutzen bieten muss (Morville, 2004). Ein dialogisches, adaptives KI-Suchtool kann diesem Anspruch gerecht werden, indem es relevante Vorschläge unterbreitet, den bisherigen Dialogverlauf einbezieht und auf individuelle Bedürfnisse eingeht. Ein solches personalisiertes, beratungsähnliches Nutzungserlebnis steigert die wahrgenommene Einfachheit der Bedienung und verbessert, gemäss dem TAM, dadurch die Einstellung zum KI-Suchtool insgesamt (Marangunić & Granić, 2015).

Intuitive und übersichtliche Gestaltung. Ein drittes zentrales Bedürfnis betrifft die intuitive und übersichtliche Gestaltung der Benutzungsoberfläche. Testpersonen wünschten sich eine klar strukturierte Seitenarchitektur mit einer gut sichtbaren Suchfunktion, bevorzugt oben rechts, sowie gut lesbaren, übersichtlich dargestellten Inhalten (vgl. Kap. 4.3.3). Diese Erkenntnisse decken sich mit den Anforderungen der ISO 9241-210 (2020), die eine menschenzentrierte Gestaltung fordert, bei der emotionale und kognitive Aspekte der Nutzung berücksichtigt werden. Klare, kohärente Visualisierungen sollen die Effizienz steigern und die Orientierung erleichtern. Im UX-Honeycomb-Modell stellt die Benutzbarkeit eine grundlegende Voraussetzung für eine positive Nutzungserfahrung dar (Morville, 2004). Darüber hinaus betont das Modell auch die Auffindbarkeit und Zugänglichkeit als zentrale Faktoren. Eine ganzheitlich positive UX entsteht nur, wenn relevante Inhalte und Funktionen schnell auffindbar sind und das System auch für verschiedene Nutzer*innengruppen, etwa Menschen mit Einschränkungen, problemlos nutzbar ist. Eine

gut auffindbare Suchfunktion, klare Navigationsstrukturen sowie barrierefreie Gestaltungselemente tragen entscheidend dazu bei, sowohl die Effizienz als auch die Zufriedenheit der Nutzenden zu fördern (Morville, 2004). Auch das TAM betont die Bedeutung der wahrgenommenen Einfachheit der Bedienung. Je intuitiver und einfacher ein System zu bedienen ist, desto geringer sind die mentalen Hürden und desto wahrscheinlicher wird es akzeptiert und genutzt (Marangunić & Granić, 2015). Ein schlichtes, aufgeräumtes Design mit konsistenter Informationshierarchie trägt somit wesentlich zur Nutzer*innenzufriedenheit und zur Nutzungsintention bei (Marangunić & Granić, 2015; Morville, 2004).

Glaubwürdige und aktuelle Ergebnisse. Schliesslich zeigte sich ein starkes Bedürfnis nach verlässlichen und aktuellen Ergebnissen. Die Nutzenden möchten den angezeigten Informationen vertrauen können, insbesondere bei sensiblen Themen wie aktuellen Reisehinweisen oder Produktspezifikationen. Dabei ist es ihnen wichtig, die Herkunft der Daten nachvollziehen zu können (vgl. Kap. 4.3.4). Eine transparente Darstellung, etwa durch nachvollziehbare Quellenangaben, stärkt das Vertrauen in das System. Die ISO 9241-210 (2020) unterstreicht, dass emotionale Aspekte wie Vertrauen und Sicherheit essenziell für die Gestaltung einer ganzheitlich positiven UX sind. Diese Anforderung spiegelt sich auch in der Facette Glaubwürdigkeit des UX-Honeycomb-Modells wider, da Vertrauen und Verlässlichkeit zu den Schlüsselfaktoren für eine durchweg positive Nutzungserfahrung zählen (Morville, 2004).

5.1.2 Identifizierte Herausforderungen

Die Ergebnisse der Tests zeigen, dass verschiedene Herausforderungen bestehen, welche die UX des KI-Suchtools beeinträchtigen können.

Schlechte Auffindbarkeit. Eine zentrale Herausforderung ist die Auffindbarkeit sowie die unklare Einführung des KI-Suchtools. Viele Testpersonen empfanden die seitlich platzierte Öffnungslasche als nicht intuitiv und wussten zunächst nicht, welche Funktionen der KI-Kompass bietet (vgl. Kap. 4.6). Die Herausforderung der Testpersonen, beim Nutzen von Links nicht mehr auf den KI-Kompass zurückkehren zu können, steht im Zusammenhang mit dem Aspekt der Auffindbarkeit des UX-Honeycomb-Modells (vgl. Kap. 4.6). Ein KI-Suchtool muss leicht auffindbar, eindeutig gekennzeichnet und selbsterklärend sein. Ist dies nicht der Fall, leidet die Gebrauchstauglichkeit (Morville, 2004; ISO 9241-210, 2020). Auch das TAM macht deutlich, wenn die erste Nutzung durch mangelnde Orientierung erschwert wird, sinkt die wahrgenommene Einfachheit der Bedienung, die Anwendung erscheint komplizierter, was das Vertrauen der Nutzenden in das KI-Suchtool beeinträchtigen kann (Marangunić & Granić, 2015).

Unpräzise Antworten. Einige Testpersonen kritisierten, dass die Antworten des KI-Kompasses zu kurz oder oberflächlich ausfielen. Insbesondere wurde bemängelt, dass wichtige Detailinformationen oder weiterführende Hinweise häufig fehlten, sodass für eine zufriedenstellende Antwort oftmals mehrere Rückfragen oder erneute Suchanfragen nötig waren (vgl. Kap. 4.6). Im Sinne des TAM kann eine solche eingeschränkte Informationsqualität die wahrgenommene Nützlichkeit senken. Ein KI-Suchtool, das gestellte Fragen nur unzureichend beantwortet, wird als weniger vorteilhaft wahrgenommen und somit seltener genutzt (Marangunić & Granić, 2015). Auch das UX-Honeycomb-Modell betont die Bedeutung der Nützlichkeit sowie der Glaubwürdigkeit. Eine mangelnde inhaltliche Tiefe untergräbt nicht nur den wahrgenommenen Nutzen, sondern kann auch das Vertrauen in die Qualität der bereitgestellten Informationen mindern (Morville, 2004).

Technische Probleme und Ablenkungen. Weiterhin wurden technische Probleme und ablenkende Elemente als störend empfunden. Gelegentliche Ausfälle des KI-Kompasses, fehlerhafte Links und lange Ladezeiten erschwerten die Nutzung und unterbrachen den Informationsfluss. Weiterlaufende Hintergrundmedien auf der Website wurden von mehreren Testpersonen als ablenkend beschrieben (vgl. Kap. 4.6). Nach ISO 9241-210 (2020) sollte die gesamte Nutzungssituation ganzheitlich optimiert werden, störende Elemente behindern dabei sowohl die Effizienz als auch die Zufriedenheit der Nutzenden. Im Kontext des TAM führt jede zusätzliche Wartezeit oder Fehlerquelle zu einem erhöhten subjektiven Aufwand und die wahrgenommene Einfachheit der Bedienung sinkt. Visuelle Störungen beeinträchtigen die Gebrauchstauglichkeit im Sinne der Facetten Attraktivität sowie Benutzbarkeit des UX-Honeycomb-Modells. Sie lenken die Aufmerksamkeit von der eigentlichen Aufgabe ab und vermindern die Konzentration (Morville, 2004).

5.1.3 Handlungsempfehlungen

Aus der Analyse der Nutzer*innenbedürfnisse und den identifizierten Herausforderungen lassen sich konkrete Handlungsempfehlungen ableiten, die sowohl auf den empirischen Erkenntnissen der Nutzer*innentests als auch auf den zugrunde liegenden Theorien basieren.

Auffindbarkeit und Erwartungssteuerung verbessern. Die Nutzer*innentests zeigen deutlich, dass Suchfunktionen von den meisten Testpersonen intuitiv im oberen rechten Bereich der Website erwartet werden (vgl. Kap. 4.2.1). Die derzeitige Lösung, den KI-Kompass in einer seitlichen Lasche zu platzieren, wurde von vielen Testpersonen zunächst übersehen oder nicht eindeutig als Suchfunktion erkannt (vgl. Kap. 4.5.4). Aus Sicht der Auffindbarkeit nach Morville (2004) empfiehlt es sich deshalb, den Zugang zum KI-Suchtool prominent oben rechts auf der Website zu platzieren und diesen Bereich klar als KI-

Suchtool zu kennzeichnen. Ein entsprechendes Icon sollte idealerweise durch einen eindeutigen Text ergänzt werden, sodass auf den ersten Blick klar ist, dass hier ein KI-Suchtool angeboten wird.

Positiv ist zu bewerten, dass die Nutzenden aktuell über verschiedene Wege, etwa die zentrale Suche, die Lasche am Seitenrand oder die klassische Lupe, letztlich auf den KI-Kompass zugreifen können. Dies fördert gemäss dem UX-Honeycomb-Modell grundsätzlich die Auffindbarkeit und Zugänglichkeit (Morville, 2004). Dennoch sollte das prominente, eindeutig benannte KI-Suchtool oben rechts als Standardzugang gestärkt werden, da sie dem gewohnten mentalen Modell entspricht und von allen Nutzer*innengruppen als erstes gesucht wird (vgl. Kap. 4.2.1).

Ein weiteres Ergebnis der Nutzer*innentests war, dass Testpersonen nach dem Anklicken von Links im KI-Kompass oft nicht mehr problemlos zum Chat zurückfanden (vgl. Kap. 4.6). Diese Herausforderung beeinträchtigt die Auffindbarkeit und kann das Nutzungserlebnis erheblich stören (Morville, 2004). Es wird daher empfohlen, nach dem Aufruf von Links eine klare Möglichkeit zur Rückkehr zur vorherigen Suchanfrage oder zum Chatverlauf einzubauen, etwa durch eine gut sichtbare Schaltfläche *Zurück zum Chatverlauf*. Auf diese Weise wird die Orientierung verbessert und die Auffindbarkeit der Suchfunktion auch nach der Navigation zu externen Inhalten sichergestellt.

Die Ergebnisse der Nutzer*innentests verdeutlichen, dass eine klare Erwartungssteuerung massgeblich zur Akzeptanz und effektiven Nutzung des KI-Suchtools beiträgt (vgl. Kap. 4.7). Ein zentrales Optimierungspotenzial liegt daher in der sofortigen und eindeutigen Kommunikation der KI-Suchtool-Funktionalität. Ein kurzes *Onboarding* könnte helfen, Erwartungen zu klären und den Funktionsumfang verständlich zu machen. So wird das Nutzungserlebnis früh im Prozess eingeordnet, was laut ISO 9241-210 (2020) die Gebrauchstauglichkeit erhöht. Zur weiteren Verbesserung der Erwartungssteuerung empfiehlt es sich, die Möglichkeiten und Grenzen des KI-Suchtools direkt auf der Startseite in Form einer kompakten Info-Box bereitzustellen. Beispielsweise könnte diese als klickbares Icon gestaltet sein, das eine strukturierte Übersicht bietet, welche über die Funktionen und Grenzen des KI-Suchtools informiert. Eine transparente Darstellung des Funktionsumfangs sowie der Limitationen, etwa der Hinweis auf den fehlenden Anspruch auf persönliche Beratung, die Unverbindlichkeit rechtlicher Aussagen oder potenzielle Einschränkungen bei Quellenangaben, beugt möglichen Missverständnissen in der Nutzung vor und beeinflusst die UX der Nutzenden (vgl. Kap. 4.7; Morville, 2004).

Dialogorientierung und Kontextualisierung stärken. Nutzende erwarten vom KI-Suchtool eine interaktive, kontextbezogene Kommunikation, die über reine Suchanfragen hinausgeht. Die Nutzer*innentests zeigen, dass dialogische Nachsteuerung als beratend wahrgenommen wird und Vertrauen sowie Orientierung stärkt (vgl. 4.3.2). Um diesem Bedürfnis gerecht zu werden, sollte der Dialogfluss optimiert, eine Verlaufsspeicherung der Chats eingeführt und das KI-Suchtool gezielt Rückfragen stellen, um präzisere Ergebnisse zu ermöglichen. Solche dialogischen Elemente erhöhen nicht nur die Relevanz der Resultate, sondern stärken laut ISO 9241-210 (2020) auch die Gebrauchstauglichkeit, da sie den Nutzungskontext explizit einbeziehen. Eine transparente Speicherung der Chatverläufe verringert die kognitive Belastung und fördert eine gebrauchstaugliche Interaktion, da laut TAM eine einfache Bedienung die Akzeptanz steigert (Marangunić & Granić, 2015).

Inhaltsqualität und Mehrwert ausbauen. Die Qualität der KI-Antworten wurde von mehreren Testpersonen als ausbaufähig eingestuft (vgl. Kap. 4.5.4). Um die wahrgenommene Nützlichkeit, einen Schlüsselfaktor im TAM (Marangunić & Granić, 2015) sowie dem UX-Honeycomb-Modell (Morville, 2004) zu steigern, sollte das Sprachmodell weiterentwickelt oder mit Zusatzinformationen angereichert werden. Dazu zählen vertiefte Informationen bei Antworten des KI-Suchtools, visuelle Inhalte wie beispielsweise Bilder von Produkten oder Reisezielen sowie strukturierte Produkt- oder Länderinformationen. Ergänzend kann die Kennzeichnung sensibler Themenbereiche, etwa durch die explizite Angabe offizieller Quellen bei Gesundheits- oder Reisehinweisen, die Glaubwürdigkeit erhöhen, ein zentrales Element im UX-Honeycomb-Modell (Morville, 2004).

Stabilität und Leistung verbessern. Verzögerte Antwortzeiten oder Systemabbrüche unterbrechen nicht nur den Informationsfluss, sondern senken nachweislich die Zufriedenheit und Akzeptanz (Marangunić & Granić, 2015). Auf Basis der Nutzer*innentests empfiehlt sich daher eine Performance-Optimierung auf Backend-Ebene, um Ladezeiten zu minimieren und Antwortprozesse zu stabilisieren (vgl. Kap. 4.7). Ergänzend sollten alle weiterführenden Links auf ihre Funktionsfähigkeit überprüft werden. Für ein ablenkungsfreies Nutzungserlebnis bietet es sich an, dynamische UI-Elemente, wie das animierte Video auf der Startseite, zu pausieren oder auszublenden, sobald der Chat geöffnet wird. Diese Reduktion auf das Wesentliche minimiert kognitive Störfaktoren und entspricht dem Prinzip der minimalistischen Gestaltung aus der ISO 9241-210 (2020).

Nutzer*innenzentrierte Weiterentwicklung. Die Implementation der Handlungsempfehlungen sollte im engen Austausch mit den Nutzenden entstehen. ISO 9241-210 (2020) fordert eine kontinuierliche Einbindung durch Tests und Feedback. Prototypen sollten frühzeitig überprüft und in kurzen Abständen weiterentwickelt werden.

Das erhöht nicht nur die Gebrauchstauglichkeit, sondern berücksichtigt auch emotionale und soziale Aspekte der Nutzung. Ein iterativer Ansatz sorgt dafür, dass bisher unbeachtete Bedürfnisse sichtbar werden und berücksichtigt werden können.

Zusammenfassende Beantwortung der Forschungsfrage. Die Nutzer*innentests belegen, dass die Testpersonen beim KI-Suchtool auf der Globetrotter-Website vor allem eine schnelle und gezielte Informationsfindung erwarten. Ebenso wünschen sie sich eine kontextsensitive, interaktive Kommunikation, eine intuitive Bedienung sowie glaubwürdige und aktuelle Ergebnisse. Die zentralen Herausforderungen bestehen in einer eingeschränkten Auffindbarkeit des KI-Suchtools, teilweise zu knappen oder wenig relevanten Antworten sowie technischen Problemen wie Ladezeiten und fehlerhaften Links. Um die UX gezielt zu verbessern, empfiehlt sich eine deutlich sichtbarere und intuitiv platzierte Integration des KI-Suchtools, eine stärkere Kontextualisierung und dialogische Führung mit gezielten Rückfragen, die Steigerung der Antwortqualität sowie eine ablenkungsfreie und übersichtliche Gestaltung der Benutzungsoberfläche. Die kontinuierliche, menschenzentrierte und iterative Weiterentwicklung des KI-Suchtools ist entscheidend, um künftige Bedürfnisse und Herausforderungen frühzeitig zu adressieren. Diese Handlungsempfehlungen basieren auf den Ergebnissen der Nutzer*innentests und werden durch die theoretischen Modelle ISO 9241-210, TAM und dem UX-Honeycomb-Modell gestützt. So kann die Akzeptanz gesteigert, das Vertrauen in das KI-Suchtool verbessert und die UX nachhaltig gefördert werden.

6 Fazit

Dieses Kapitel fasst die wesentlichen Limitationen der Arbeit zusammen und reflektiert kritisch das Forschungsdesign sowie die Aussagekraft der erhobenen Daten. Anschliessend wird ein Ausblick auf weiterführende Forschungsfragen und zukünftige Entwicklungsmöglichkeiten gegeben. Im Zentrum steht dabei, inwiefern die Ergebnisse und Handlungsempfehlungen bei zukünftigen Integrationen auf anderen Websites übertragbar sind und welche Aspekte in der weiteren Forschung besondere Beachtung verdienen.

6.1 Limitationen und Reflexion des Forschungsdesigns

Die vorliegende Arbeit weist mehrere Limitationen auf, die die Aussagekraft der Ergebnisse einschränken können. Die Stichprobe umfasste lediglich sieben Testpersonen, die aus dem erweiterten Umfeld des Autors rekrutiert wurden. Trotz der bewussten Auswahl ohne enge persönliche Beziehungen lässt sich eine gewisse Verzerrung durch diese Nähe nicht vollständig ausschliessen. Zudem wurden die Nutzer*innentests ausschliesslich remote durchgeführt, wodurch die direkte Beobachtung nonverbaler Kommunikation und des Nutzungskontexts nur eingeschränkt möglich war.

Eine zentrale methodische Limitation ist der bewusste Verzicht auf quantitative Erhebungen. Dieser Ansatz war insbesondere deshalb sinnvoll, weil der Fokus der Arbeit auf UX und nicht auf Gebrauchstauglichkeit lag. Qualitative Methoden eignen sich besonders, um subjektive Wahrnehmungen, individuelle Nutzungserfahrungen und vielschichtige Nutzer*innenbedürfnisse zu erfassen (Helfferich, 2011). Quantitative Instrumente wie der UEQ- oder AttrakDiff-Fragebogen liefern zwar wertvolle Vergleichswerte bei grossen Stichproben, hätten in diesem Kontext mit nur sieben Testpersonen jedoch keinen zusätzlichen Mehrwert gebracht. Bei grösseren und repräsentativeren Stichproben könnten derartige standardisierte Fragebögen die Ergebnisse sinnvoll ergänzen.

Die Erfahrung des Autors in der Durchführung von Nutzer*innentests war zum Zeitpunkt der Erhebung begrenzt, was sich potenziell auf die Qualität der Datenerhebung ausgewirkt haben könnte. Eine weitere methodische Herausforderung ergab sich aus der Reihenfolge der Befragung. Um unverfälschte Einstellungen bezüglich KI, Erwartungen zu Suchfunktionen und mentale Modelle zu erheben, wurden diese vor der Interaktion mit dem KI-Suchtool abgefragt. Dadurch sollte verhindert werden, dass die Nutzungserfahrung selbst diese Einstellungen beeinflusst. Gleichzeitig besteht bei diesem Vorgehen das Risiko, dass die Vorbefragung bereits eine Sensibilisierung für bestimmte Themenbereiche, wie etwa *KI*, bewirkt und dadurch die spätere Wahrnehmung im Nutzer*innentest beeinflusst. Diese potenzielle Beeinflussung muss bei der Interpretation der Ergebnisse berücksichtigt werden.

Ein weiterer wichtiger Aspekt betrifft die Kontextgebundenheit der Untersuchungsergebnisse. Die gewonnenen Erkenntnisse beziehen sich spezifisch auf die Integration und Nutzung des KI-Suchtools auf der Globetrotter-Website und sind eng mit diesem Anwendungskontext verknüpft. Eine direkte Übertragung der Ergebnisse auf andere Plattformen ist daher nur eingeschränkt möglich. Im Ausblick dieser Arbeit wird jedoch diskutiert, inwiefern zentrale Erkenntnisse und Handlungsempfehlungen auch auf vergleichbare Integrationen von KI-Suchtools in anderen digitalen Umgebungen übertragbar und anpassbar sind (vgl. Kap. 6.2).

Insgesamt ermöglichte das gewählte Forschungsdesign, insbesondere die Kombination aus qualitativen Remote-Nutzer*innentests, Think-Aloud-Methode und strukturierten Befragungen, einen tiefen Einblick in individuelle Nutzungserfahrungen, subjektive Wahrnehmungen und konkrete Nutzungsherausforderungen. Somit konnte die Fragestellung vollumfänglich beantwortet werden. Die Generalisierbarkeit der Ergebnisse bleibt jedoch durch die in diesem Kapitel genannten Limitationen begrenzt.

6.2 Ausblick und Forschungsimplicationen

Die Erkenntnisse dieser Arbeit bieten eine wertvolle Grundlage für die Weiterentwicklung und Integration von KI-Suchtools, sowohl auf der Globetrotter-Website als auch in anderen digitalen Kontexten. Besonders deutlich wurde, dass Nutzende das KI-Suchtool eng mit der Marke *Globetrotter* verbinden (vgl. Kap. 4.5.6). Daraus leitet sich für die Praxis ab, dass eine konsequente Anpassung an das Corporate Design sowie eine stimmige Einbettung in die Markenidentität unerlässlich sind, um Vertrauen und Glaubwürdigkeit nachhaltig zu sichern.

Die zentralen Handlungsempfehlungen dieser Arbeit lassen sich als praxisnahe Leitlinien grundsätzlich auf verschiedene Branchen und Anwendungsfälle übertragen. Das in dieser Arbeit eingesetzte UX-Honeycomb-Modell hat sich dabei als wirkungsvolles Analyseinstrument erwiesen, um die UX systematisch und ganzheitlich zu erfassen und weiterzuentwickeln. Für eine erfolgreiche Implementierung von KI-Suchtools müssen alle sieben Dimensionen konsequent berücksichtigt werden, um die Akzeptanz und Zufriedenheit der Zielgruppen zu maximieren.

Für die Forschung eröffnen sich mehrere vielversprechende Ansatzpunkte. Besonders lohnend erscheint ein vertiefender Vergleich zwischen technikaffinen und weniger technikaffinen Nutzer*innengruppen. Zwar wurde diese Differenz in der vorliegenden Arbeit im Sampling berücksichtigt, jedoch nicht systematisch analysiert. Zukünftige Studien könnten hier wertvolle Erkenntnisse zur differenzierten Gestaltung und Integration von KI-Suchtools liefern, etwa hinsichtlich der Ausgestaltung von Hilfestellungen, dem Einsatz einer Infobox oder der Anpassung von UI-Komplexität an verschiedene Zielgruppen.

Ein weiterer Forschungsbedarf besteht in der Analyse der Langzeitnutzung des KI-Suchtools. Die vorliegende qualitative Methodik ermöglicht vor allem Einblicke in die ersten Eindrücke und subjektiven Bewertungen der Nutzenden. Um langfristige Nutzungsmuster, die Entwicklung der Akzeptanz sowie Veränderungen im Suchverhalten besser zu verstehen, sind ergänzende quantitative Analysen, etwa über Nutzungsdaten oder längsschnittliche Erhebungen, empfehlenswert.

Darüber hinaus sollten die Wechselwirkungen zwischen KI-Suchtools und klassischen Suchfunktionen genauer untersucht werden. Gerade im Hinblick auf parallele Nutzung, unterschiedliche Aufgabenstellungen und individuelle Präferenzen können sich wichtige Ansatzpunkte für eine optimierte Integration ergeben.

6.3 Schlusswort

Die in dieser Arbeit gewonnenen Erkenntnisse liefern einen praxisnahen Orientierungsrahmen für die Integration und Optimierung von KI-Suchtools und eröffnen zahlreiche Perspektiven für weiterführende Forschung. Durch die empirisch fundierte Analyse der Nutzer*innenbedürfnisse und Herausforderungen wurde ein wichtiger Grundstein gelegt, um digitale Sucherlebnisse künftig noch gezielter an den tatsächlichen Erwartungen der Nutzenden auszurichten. So kann das Potenzial von KI-Suchtools nicht nur innovativ, sondern auch verantwortungsvoll und menschenzentriert erschlossen werden.

Literaturverzeichnis

- Brush, A. J. B., Ames, M. & Davis, J. (2004). A comparison of synchronous remote and local usability studies for an expert interface. *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (S. 1179–1182). Gehalten auf der CHI04: CHI 2004 Conference on Human Factors in Computing Systems, Vienna Austria: ACM. <https://doi.org/10.1145/985921.986018>
- Corritore, C. L., Kracher, B. & Wiedenbeck, S. (2003). On-line trust: concepts, evolving themes, a model (Trust and Technology). *International Journal of Human-Computer Studies*, 58(6), 737–758. [https://doi.org/10.1016/S1071-5819\(03\)00041-7](https://doi.org/10.1016/S1071-5819(03)00041-7)
- Davis, F. D. (1985). *A technology acceptance model for empirically testing new end-user information systems: theory and results*. Thesis. Massachusetts Institute of Technology. Zugriff am 26.5.2025. Verfügbar unter: <https://dspace.mit.edu/handle/1721.1/15192>
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319–340. Management Information Systems Research Center, University of Minnesota. <https://doi.org/10.2307/249008>
- Davis, F. D. & Venkatesh, V. (1996). A critical assessment of potential measurement biases in the technology acceptance model: three experiments. *International Journal of Human-Computer Studies*, 45(1), 19–45. <https://doi.org/10.1006/ijhc.1996.0040>
- De Blecker, I. & Okoroji, R. (2018). *Remote Usability Testing*. Birmingham - Mumbai: Packt Publishing. Zugriff am 11.3.2025. Verfügbar unter: <https://learning.oreilly.com/library/view/remote-usability-testing/9781788999045/>
- Diefenbach, S. & Hassenzahl, M. (2017). Ein erstes Modell erlebnisbezogener Produktqualitäten: hedonisch versus pragmatisch. In S. Diefenbach & M. Hassenzahl (Hrsg.), *Psychologie in der nutzerzentrierten Produktgestaltung: Mensch-Technik-Interaktion-Erlebnis* (S. 57–88). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-662-53026-9_3
- Hassenzahl, M. (2004). The Interplay of Beauty, Goodness, and Usability in Interactive Products. *Human-Computer Interaction*, 19(4), 319–349. Taylor & Francis. https://doi.org/10.1207/s15327051hci1904_2
- Helfferich, C. (2011). *Die Qualität qualitativer Daten*. Wiesbaden: VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-92076-4>
- Kuckartz, U. & Rädiker, S. (2020). *Fokussierte Interviewanalyse mit MAXQDA: Schritt für Schritt*. Wiesbaden: Springer Fachmedien. <https://doi.org/10.1007/978-3-658-31468-2>

- Marangunić, N. & Granić, A. (2015). Technology acceptance model: a literature review from 1986 to 2013. *Universal Access in the Information Society*, 14(1), 81–95.
<https://doi.org/10.1007/s10209-014-0348-1>
- Morville, P. (2004). *User Experience Design*. Zugriff am 8.2.2025. Verfügbar unter:
https://semanticstudios.com/user_experience_design/
- Morville, P. (2005). *Ambient Findability*. Sebastopol: O'Reilly Media, Inc. Zugriff am 8.2.2025. Verfügbar unter: <https://learning.oreilly.com/library/view/ambient-findability/0596007655/>
- Moser, C. (2012). *User Experience Design: Mit erlebniszentrierter Softwareentwicklung zu Produkten, die begeistern* (X.media.press). Berlin, Heidelberg: Springer.
<https://doi.org/10.1007/978-3-642-13363-3>
- Nielsen, J. (1994). *Usability Engineering*. San Francisco: Morgan Kaufmann.
- Nielsen, J. (2012). Thinking Aloud: The #1 Usability Tool. *Nielsen Norman Group*. Zugriff am 17.3.2025. Verfügbar unter: <https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/>
- Norman, D. A. (2013). *The design of everyday things* (Rev. and expanded edition.). Cambridge (Mass.): MIT press.
- Petrie, H. & Kheir, O. (2007). The relationship between accessibility and usability of websites. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 397–406). Gehalten auf der CHI07: CHI Conference on Human Factors in Computing Systems, San Jose California USA: ACM.
<https://doi.org/10.1145/1240624.1240688>
- Rädiker, S. & Kuckartz, U. (2019). *Analyse qualitativer Daten mit MAXQDA: Text, Audio und Video*. Wiesbaden: Springer Fachmedien. <https://doi.org/10.1007/978-3-658-22095-2>
- Richter, M. & Flückiger, M. D. (2016). *Usability und UX kompakt* (IT kompakt). Berlin, Heidelberg: Springer. <https://doi.org/10.1007/978-3-662-49828-6>
- Schweizerische Normen-Vereinigung. (2020). *Ergonomie der Mensch-System-Interaktion - Teil 210: Prozess zur Gestaltung gebrauchstauglicher interaktiver Systeme (ISO 9241-210:2019)*. Winterthur. Zugriff am 1.3.2025. Verfügbar unter:
<https://viewer.snv.ch/product/270226?langUI=de&filePath=100798557.pdf&fileType=Pdf>
- Spatharioti, S. E., Rothschild, D. M., Goldstein, D. G. & Hofman, J. M. (2023). Comparing Traditional and LLM-based Search for Consumer Choice: A Randomized Experiment. arXiv. <https://doi.org/10.48550/ARXIV.2307.03744>

- Thompson, K. E., Rozanski, E. P. & Haake, A. R. (2004). Here, there, anywhere: remote usability testing that works. *Proceedings of the 5th conference on Information technology education* (S. 132–137). Gehalten auf der SIGITE04: ACM Special Interest Group for Information Technology Education Conference 2004, Salt Lake City UT USA: ACM. <https://doi.org/10.1145/1029533.1029567>
- Tractinsky, N., Katz, A. S. & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers*, 13(2), 127–145. [https://doi.org/10.1016/S0953-5438\(00\)00031-X](https://doi.org/10.1016/S0953-5438(00)00031-X)
- Tuch, A. N., Presslauer, E. E., Stöcklin, M., Opwis, K. & Bargas-Avila, J. A. (2012). The role of visual complexity and prototypicality regarding first impression of websites: Working towards understanding aesthetic judgments. *International Journal of Human-Computer Studies*, 70(11), 794–811. <https://doi.org/10.1016/j.ijhcs.2012.06.003>
- Xiong, H., Bian, J., Li, Y., Li, X., Du, M., Wang, S. et al. (2024). When Search Engine Services Meet Large Language Models: Visions and Challenges. *IEEE Transactions on Services Computing*, 17(6), 4558–4577. <https://doi.org/10.1109/TSC.2024.3451185>

Abbildungsverzeichnis

Abbildung 1. Die Phasen der menschenzentrierten Gestaltung, von der Kontextanalyse bis zur Evaluation, sind in einem wechselseitig abhängigen und iterativen Zyklus organisiert (ISO 9241-210, 2020, S. 21).....	5
Abbildung 2. Das TAM: Die wahrgenommene Nützlichkeit und die wahrgenommene Einfachheit der Bedienung beeinflussen die Benutzungsintention und somit die tatsächliche Benutzung (Davis & Venkatesh, 1996, S. 2).....	7
Abbildung 3. Die sieben Facetten des UX-Honeycomb-Modells erinnern an eine Bienenwabe, um die enge Verzahnung der UX-Faktoren zu symbolisieren. Eigene Darstellung nach Morville (2004)	9
Abbildung 4. SPSS-Prinzip nach Helfferich (2011), eigene Darstellung. Das Prinzip visualisiert den Prozess der qualitativen Interviewleitfadenerstellung durch die vier Schritte Sammeln, Prüfen, Sortieren und Subsumieren.....	14
Abbildung 5. Der Datenschutz-Hyperlink wurde von den Testpersonen nicht selbstständig genutzt. Im Nachhinein bewerteten sie den Link jedoch als äusserst positiv (B01 - B07). Verfügbar unter: https://www.globetrotter.ch/	26
Abbildung 6. B04 merkte an, dass die rot markierte Suchleiste des KI-Kompasses zu weit unten ist und wünschte sich eine prominentere Platzierung. Screenshot aus Nutzer*innentest B04.....	29
Abbildung 7. B01 empfand die Informationsseite über den KI-Kompass als visuell erdrückend. Verfügbar unter: https://www.globetrotter.ch/ki-kompass-wie-funktioniert	29
Abbildung 8. Der KI-Kompass gibt auf die Frage nach Beratung für Mallorca an, ein Reisebüro aufzusuchen und verweist auf Mallorca Tourismus. Screenshot aus Nutzer*innentest B05.....	34
Abbildung 9. Testpersonen haben erwartet, mithilfe des rot markierten Zurück-Buttons wieder auf den Chat mit dem KI-Kompass zu gelangen. Seite verfügbar unter: https://www.globetrotter.ch/reise/mallorca-mit-dem-mietwagen-unterwegs	35
Abbildung 10. Der KI-Kompass reagiert nicht auf die Eingaben der Testperson. Screenshot aus dem Nutzer*innentest B04	36
Abbildung 11. Durch einen Link in einer Antwort bekam B02 eine Fehlermeldung. Screenshot aus Nutzer*innentest B02	36

Tabellenverzeichnis

Tabelle 1 Sampling Nutzer*innentests	17
--	----

Abkürzungsverzeichnis

KI	<i>künstliche Intelligenz</i>
KI-Suchtool.....	<i>KI-gestütztes Suchtool</i>
TAM	<i>Technologieakzeptanzmodell</i>
UI	<i>User Interface</i>
UX.....	<i>User Experience</i>

Hilfsmittelverzeichnis mit Verwendungszweck

KI-Assistenzsystem	Teile / Stelle(n) in der Arbeit	Einsatz
ChatGPT	Abstract	Kürzen Abstract
ChatGPT	Kapitel 1 - 6	Überprüfung Rechtschreibfehler / Umschreiben von Sätzen
Elicit / Connected Papers	Kapitel 2 - Theorie	Suche nach geeigneten Studien
ChatGPT	Kapitel 3 - Methodik	Exploration Interviewfragen / Hypothesen / Testaufgaben