

Hochschule für angewandte Wissenschaften Ansbach

Fakultät Medien

Bachelorarbeit zum Thema

**Wenn Algorithmen gestalten: In einer Vergleichsstudie treten eine KI-generierte Applikation und eine manuell gestaltete Applikation in Bezug auf die Benutzererfahrung und die Benutzerfreundlichkeit gegeneinander an.**

Zur Erlangung des Akademischen Grades

Bachelor of Arts

Vorgelegt von:

**Marie Sophie Seitz**

Matrikelnummer 00160069

Fachsemester 10

Studiengang:	Visualisierung und Interaktion in digitalen Medien
Erstgutachter:	Prof. Dr.-Ing. Sebastian Stadler
Zweitgutachter	Prof. Christian Barta
Abgabedatum:	05.05.2025



# Inhalt

<b>Eigenständigkeitserklärung .....</b>	I
<b>Abkürzungsverzeichnis .....</b>	II
<b>Abbildungsverzeichnis .....</b>	III
<b>Tabellenverzeichnis .....</b>	VIII
1. Einleitung .....	1
1.1 Forschungsfrage und Ziel .....	2
1.2 Aufbau der Arbeit .....	2
2. Theorie .....	3
2.1 Grundlagen Webdesign .....	3
2.1.1 User Experience.....	3
2.1.2 Usability.....	5
2.1.3 User Interface.....	6
2.1.4 Gesetze für Webdesign.....	7
2.1.5 Der Human-Centered Design-Prozess .....	17
2.2 Grundlagen KI .....	22
2.2.1 Basiswissen KI .....	22
2.2.2 KI im Webdesign .....	24
2.3 Grundlagen der Datenanalyse .....	25
2.3.1 Mann-Whitney-U-Test.....	25
2.3.2 Deskriptive Statistik.....	30
3. Methodik .....	32
3.1 Erstellung der manuell gestalteten Applikation .....	32
3.1.1 Verstehen und Beschreiben des Nutzungskontextes .....	32
3.1.1.1 Externes Benchmarking .....	32
3.1.1.2 Tests bestehender Anwendungen .....	45
3.1.1.3 Feedback Capture Grid.....	45
3.1.1.4 Customer Journey Map.....	51
3.1.2 Nutzungsanforderungen spezifizieren .....	54
3.1.2.1 Personas .....	54

**Hinweis:** Aus Gründen der besseren Lesbarkeit wird auf die gleichzeitige Verwendung der Sprachformen männlich, weiblich und divers (m/w/d) verzichtet. Sämtliche Personenbezeichnungen gelten gleichermaßen für alle Geschlechter.

3.1.2.2	User Stories.....	57
3.1.2.3	How might we...?.....	59
3.1.3	Lösungsentwicklung auf Basis der Nutzungsanforderungen.....	60
3.1.3.1	Brainstorming .....	60
3.1.3.2	Hybrides Card Sorting.....	62
3.1.3.3	Future-State Journey Map .....	68
3.1.3.4	Informationsarchitektur.....	70
3.1.3.5	Wireframing .....	71
3.1.3.6	Styleguide.....	73
3.1.3.7	Low-Fidelity-Prototype .....	80
3.1.3.8	High-Fidelity-Prototype 1.....	83
3.1.3.9	High-Fidelity-Prototype 2.....	86
3.2	Erstellung der KI-generierten Applikation .....	87
3.2.1	Programmauswahl .....	87
3.2.2	Erstellung .....	88
3.2.2.1	Kriterien für die Erstellung .....	88
3.2.2.2	Vorbereitung der Erstellung .....	88
3.2.2.3	Ablauf der Erstellung.....	89
3.2.2.4	Beobachtungen während der Erstellung.....	89
3.2.3	Fazit.....	97
3.3	Testen der Lösungen gegen die Nutzungsanforderungen .....	98
3.3.1	Usability Test.....	98
3.3.2	Umfragen.....	99
4.	Ergebnisse des Usability-Tests .....	101
4.1	Ergebnisse des Mann-Whitney-U-Test .....	101
4.1.1	Statistisch signifikante Datensätze .....	101
4.1.2	Statistisch nicht signifikante Datensätze.....	102
4.2	Ergebnisse der deskriptiven Statistik .....	103
4.3	Auswertung der Mitschriften und der Umfragen .....	104
4.3.1	Feedback Capture Grid der KI-generierten Applikation .....	104
4.3.2	Feedback Capture Grid der manuell gestalteten Applikation .....	111

4.3.3	Emotionen während der Tests der KI-generierten Applikation .....	119
4.3.4	Emotionen während der Tests der manuell gestalteten Applikation.....	121
4.3.5	Verhalten während der Tests der KI-generierten Applikation .....	123
4.3.6	Verhalten während der Tests der manuell gestalteten Applikation.....	125
4.4	Abschluss der Aufgaben .....	126
5.	Interpretation der Ergebnisse .....	127
5.1	Statistisch nicht signifikante Datensätze.....	127
5.2	Statistisch signifikante Datensätze.....	132
6.	Diskussion der Ergebnisse.....	157
7.	Fazit.....	160
7.1	Zusammenfassung.....	160
7.2	Zukunftsansicht.....	161
7.3	Bewertung der Arbeit .....	161
8.	Verwendete Tools.....	162
9.	Referenzen .....	163
	Anhangsverzeichnis .....	169

# Eigenständigkeitserklärung

"Ich versichere, dass ich die Arbeit selbständig angefertigt, nicht anderweitig für Prüfungszwecke vorgelegt, alle benützten Quellen und Hilfsmittel angegeben sowie wörtliche und sinngemäße Zitate gekennzeichnet habe."

Marie Sophie Seitz	00160069	<u>Marie Seitz</u>
Name	Matrikelnummer	Unterschrift

Einer möglichen Präsentation meiner Arbeit durch die HS-Ansbach stimme(n) ich/wir zu. Die Verwendung von urheberrechtlich geschützten Bestandteilen ist nicht zulässig, außer der Rechteinhaber stimmt nachweislich der Verwendung zu (Nachweis beifügen).

## **Abkürzungsverzeichnis**

KI Künstliche Intelligenz

UI User Interface

UX User Experience

UEQ User Experience Questionnaire

IQR Interquartilsabstand

# Abbildungsverzeichnis

Abbildung 1. Unterschied zwischen User Experience und Usability (Jacobsen & Meyer, 2022, S. 33)	4
Abbildung 2. Nielsen's 10 Usability Heuristiken (Nielsen, 1994), Grafik: Agente (2019)	6
Abbildung 3. Ontologisches Designdiagramm von Gui Bonsiepe von 1996 (Thesmann, 2009, S. 7)	7
Abbildung 4. Gesetz der Nähe am Beispiel des Google-Suchergebnis (Hahn, 2020, S. 223)	8
Abbildung 5. Gesetz der Geschlossenheit am Beispiel einer Anmeldeseite (Hahn, 2020, S. 225)	9
Abbildung 6. Nicht berücksichtigtes Gesetz der Geschlossenheit (Hahn, 2020, S. 226)	9
Abbildung 7. Gesetz der Ähnlichkeit am Beispiel von toom.de (Hahn, 2020, S. 227)	10
Abbildung 8. Gesetz der Erfahrung anhand von Symbolen (Hahn, 2020, S. 229)	11
Abbildung 9. Jakobs Law anhand von Etsy erklärt (Quelle: Etsy, 2025)	12
Abbildung 10. Gesetz der Kontinuität am Beispiel von Seitenzahlen (Hahn, 2020, S.)	12
Abbildung 11. Touch-Genauigkeit bei Smartphones (Yablonski, 2024, S. 36)	13
Abbildung 12. Fitts' Gesetz anhand eines Formularfeldes (Yablonski, 2024, S. 37)	13
Abbildung 13. Millers Gesetz: Beispiel einer Textwüste ( <i>Chapter 4. Miller's Law</i> , n.d.)	14
Abbildung 14. Millers Gesetz: Verbesserte Textwüste ( <i>Chapter 4. Miller's Law</i> , n.d.)	15
Abbildung 15. Hicks Gesetz anhand der Google-Suche (Quelle: Google, 2025)	16
Abbildung 16. Phasen des Human-Centered Design-Prozess (Heimgärtner, 2017, S. 106)	17
Abbildung 17. Folgen von Künstlicher Intelligenz für die Arbeitsbedingungen in Unternehmen (Zanker et al., 2019, S. 28)	23
Abbildung 18. Ausschnitt der Tabelle zur Verteilungsfunktion der Standardnormalverteilung (Bortz & Schuster, 2010, S. 587)	28
Abbildung 19. Ausschnitt der Tabelle zur Verteilungsfunktion der Standardnormalverteilung für einen hohen z-Wert (Uni Hamburg, n.d.)	29
Abbildung 20. Veranschaulichung der Inhalte des Boxplot-Diagramms	31
Abbildung 21. Benchmarking – Kriterien der Vergleichspartner	33
Abbildung 22. Benchmarking – Wurde das Problem bereits untersucht?	33
Abbildung 23. Benchmarking – Datengewinnung für die Website von Tee Kontor Kiel	34
Abbildung 24. Benchmarking – Datengewinnung für die Website von Königsmann Teeladen	34
Abbildung 25. Benchmarking – Datengewinnung für die Website von Tee Gschwendner	35
Abbildung 26. Benchmarking – Datengewinnung für die Website von The English Tearoom	35
Abbildung 27. Benchmarking – Zielgruppe	36
Abbildung 28. Benchmarking – Ermittelte Stärken von Tee Kontor Kiel	37
Abbildung 29. Benchmarking – Ermittelte Stärken von Tee Gschwendner	38
Abbildung 30. Benchmarking – Ermittelte Stärken von The English Tearoom	38
Abbildung 31. Benchmarking – Ermittelte Stärken von Königsmann Teeladen	39
Abbildung 32. Benchmarking – Ermittelte Schwächen von Tee Kontor Kiel	39
Abbildung 33. Benchmarking – Ermittelte Schwächen von Tee Gschwendner	40
Abbildung 34. Benchmarking – Ermittelte Schwächen von The English Tearoom	41
Abbildung 35. Benchmarking – Ermittelte Schwächen von Königsmann Teeladen	41

Abbildung 36. Benchmarking – Ermittelte markante Unterschiede.....	42
Abbildung 37. Benchmarking – Lösen die Produkte das Problem? .....	43
Abbildung 38. Benchmarking – Was muss in der eigenen Anwendung umgesetzt werden? .....	44
Abbildung 39. Benchmarking – Zusammenfassung der Erkenntnisse der Methode .....	45
Abbildung 40. Feedback Capture Grid für Tee Kontor Kiel – What worked.....	46
Abbildung 41. Feedback Capture Grid für Tee Kontor Kiel – What could be improved.....	47
Abbildung 42. Feedback Capture Grid für Tee Kontor Kiel – Ideas .....	48
Abbildung 43. Feedback Capture Grid für The English Tearoom – What worked .....	49
Abbildung 44. Feedback Capture Grid für The English Tearoom – What could be improved .....	50
Abbildung 45. Feedback Capture Grid für The English Tearoom – Ideas .....	50
Abbildung 46. Customer Journey Map für Tee Kontor Kiel – Erfahrener Teetrinker .....	51
Abbildung 47. Customer Journey Map für Tee Kontor Kiel – Unerfahrener Teetrinker .....	52
Abbildung 48. Customer Journey Map für The English Tearoom – Erfahrener Teetrinker.....	53
Abbildung 49. Customer Journey Map für The English Tearoom – Unerfahrener Teetrinker .....	54
Abbildung 50. Persona des erfahrenen Teetrinkers.....	55
Abbildung 51. Persona des etwas erfahrenen Teetrinkers .....	56
Abbildung 52. Persona des unerfahrenen Teetrinkers.....	57
Abbildung 53. Ideen aus dem Brainstorming unter dem Oberbegriff Inhalt .....	61
Abbildung 54. Ideen aus dem Brainstorming unter dem Oberbegriff Gestaltung .....	62
Abbildung 55. Hybrides Card Sorting – Oberbegriff Home .....	63
Abbildung 56. Hybrides Card Sorting – Oberbegriff Tee.....	63
Abbildung 57. Hybrides Card Sorting – Oberbegriff Matcha .....	64
Abbildung 58. Hybrides Card Sorting – Oberbegriff Zubehör .....	65
Abbildung 59. Hybrides Card Sorting – Oberbegriff Köstlichkeiten.....	66
Abbildung 60. Hybrides Card Sorting – Oberbegriff Laden.....	67
Abbildung 61. Hybrides Card Sorting – Oberbegriff Gutscheine.....	67
Abbildung 62. Hybrides Card Sorting – Oberbegriff Header Allgemein .....	68
Abbildung 63. Future Journey Map – Erfahrener Teetrinker.....	69
Abbildung 64. Future Journey Map – Unerfahrener Teetrinker.....	70
Abbildung 65. Erster Entwurf der Seite „Köstlichkeiten“ .....	71
Abbildung 66. Zweiter Entwurf der Seite „Köstlichkeiten“ .....	72
Abbildung 67. Entwurf des Online-Shops für Tee .....	73
Abbildung 68. Farben der Hintergründe .....	74
Abbildung 69. Produkt ohne und mit Hover-Effekt .....	74
Abbildung 70. Farben der Buttons.....	75
Abbildung 71. Button ohne und mit Hover-Effekt .....	75
Abbildung 72. Produkt ohne und mit Hover-Effekt bei dem Button .....	75
Abbildung 73. Warenkorb Produkt ohne und mit Hover-Effekt .....	76
Abbildung 74. Farben der Schriften.....	76
Abbildung 75. Farbe der Icons .....	76

Abbildung 76. Icons im Header .....	77
Abbildung 77. Icons unter Home .....	77
Abbildung 78. Typographie.....	77
Abbildung 79. Erste Skizzen des Logos .....	78
Abbildung 80. Logo des High Fidelity Prototyp 1 .....	79
Abbildung 81. Finales Logo .....	79
Abbildung 82. Header .....	79
Abbildung 83. Footer .....	80
Abbildung 84. Feedback Capture Grid für den Low-Fidelity-Prototype – What worked.....	81
Abbildung 85. Feedback Capture Grid für den Low-Fidelity-Prototype – What could be improved.....	82
Abbildung 86. Feedback Capture Grid für den Low-Fidelity-Prototype – Ideas.....	82
Abbildung 87. Feedback Capture Grid für den High-Fidelity-Prototype – What worked.....	84
Abbildung 88. Feedback Capture Grid für den High-Fidelity-Prototype – What could be improved.....	85
Abbildung 89. Feedback Capture Grid für den High-Fidelity-Prototype – Ideas .....	86
Abbildung 90. Abzuarbeitende Übersicht bei der Erstellung der KI-Applikation .....	90
Abbildung 91. Erstes Design der KI-Applikation.....	91
Abbildung 92. Weitere Ansicht des ersten Designs der KI-Applikation.....	91
Abbildung 93. Finales Design der KI-Applikation .....	93
Abbildung 94. Zweite Ansicht des finalen Designs der KI-Applikation .....	93
Abbildung 95. Banner zur weiteren Unterscheidung der Teesorten .....	95
Abbildung 96. Finaler Online-Shop der KI-Applikation .....	95
Abbildung 97. Erstellter Abschnitt „Unser Team“ .....	96
Abbildung 98. Öffnungszeiten der KI-generierten Applikation .....	96
Abbildung 99. Beispielfrage des UEQ .....	100
Abbildung 100. Beispielfrage für die zufällige Reihenfolge der Begriffe des UEQ.....	100
Abbildung 101. Feedback Capture Grid für die Mitschriften der KI-generierten Applikation .....	105
Abbildung 102. Feedback Capture Grid für die Mitschriften der KI-generierten Applikation .....	106
Abbildung 103. Feedback Capture Grid für die Mitschriften der KI-generierten Applikation .....	106
Abbildung 104. Feedback Capture Grid für die Mitschriften der KI-generierten Applikation .....	107
Abbildung 105. Feedback Capture Grid für die Umfrage der KI-generierten Applikation .....	108
Abbildung 106. Feedback Capture Grid für die Umfrage der KI-generierten Applikation .....	109
Abbildung 107. Feedback Capture Grid für die Umfrage der KI-generierten Applikation .....	109
Abbildung 108. Feedback Capture Grid für die Umfrage der KI-generierten Applikation .....	110
Abbildung 109. Feedback Capture Grid für die Mitschrift der manuell gestalteten Applikation .....	111
Abbildung 110. Feedback Capture Grid für die Mitschrift der manuell gestalteten Applikation .....	112
Abbildung 111. Feedback Capture Grid für die Mitschrift der manuell gestalteten Applikation .....	113
Abbildung 112. Feedback Capture Grid für die Mitschrift der manuell gestalteten Applikation .....	114
Abbildung 113. Feedback Capture Grid für die Mitschrift der manuell gestalteten Applikation .....	114
Abbildung 114. Feedback Capture Grid für die Umfrage der manuell gestalteten Applikation.....	115
Abbildung 115. Feedback Capture Grid für die Umfrage der manuell gestalteten Applikation.....	116

Abbildung 116. Feedback Capture Grid für die Umfrage der manuell gestalteten Applikation.....	117
Abbildung 117. Feedback Capture Grid für die Umfrage der manuell gestalteten Applikation.....	118
Abbildung 118. Feedback Capture Grid für die Umfrage der manuell gestalteten Applikation.....	118
Abbildung 119. Positive Emotionen während des Nutzens der KI-generierten Applikation.....	119
Abbildung 120. Negative Emotionen während des Nutzens der KI-generierten Applikation .....	120
Abbildung 121. Weitere Negative Emotionen während des Nutzens der KI-generierten Applikation.	120
Abbildung 122. Neutrale Emotionen während des Nutzens der KI-generierten Applikation.....	121
Abbildung 123. Positive Emotionen während des Nutzens der manuell gestalteten Applikation .....	122
Abbildung 124. Negative Emotionen während des Nutzens der manuell gestalteten Applikation.....	123
Abbildung 125. Positives Verhalten während des Nutzens der KI-generierten Applikation.....	123
Abbildung 126. Negatives Verhalten während des Nutzens der KI-generierten Applikation .....	124
Abbildung 127. Weiteres negatives Verhalten während des Nutzens der KI-generierten Applikation	125
Abbildung 128. Positives Verhalten während des Nutzens der manuell gestalteten Applikation .....	125
Abbildung 129. Verhalten während des Nutzens der manuell gestalteten Applikation.....	126
Abbildung 130. Boxplot-Diagramm für die Bewertung der Schnelligkeit.....	127
Abbildung 131. Boxplot-Diagramm für die Bewertung der Erlernbarkeit .....	128
Abbildung 132. Boxplot-Diagramm für die Bewertung der Stimulierung .....	129
Abbildung 133. Boxplot-Diagramm für die Bewertung der Kreativität.....	130
Abbildung 134. Boxplot-Diagramm für die Bewertung der Sicherheit .....	131
Abbildung 135. Boxplot-Diagramm für die Bewertung der Übersichtlichkeit.....	132
Abbildung 136. Boxplot-Diagramm für die Bewertung der Unterstützung .....	133
Abbildung 137. Boxplot-Diagramm für die Bewertung der Verständlichkeit .....	134
Abbildung 138. Boxplot-Diagramm für die Bewertung der Effizienz .....	135
Abbildung 139. Boxplot-Diagramm für die Bewertung des Gefallens .....	136
Abbildung 140. Boxplot-Diagramm für die Bewertung der Erfreulichkeit .....	137
Abbildung 141. Boxplot-Diagramm für die Bewertung der Anziehung .....	138
Abbildung 142. Boxplot-Diagramm für die Bewertung des Interesses.....	139
Abbildung 143. Boxplot-Diagramm für die Bewertung der Navigation .....	140
Abbildung 144. Boxplot-Diagramm für die Bewertung des Wohlbefindens .....	141
Abbildung 145. Boxplot-Diagramm für die Bewertung der Spannung.....	142
Abbildung 146. Boxplot-Diagramm für die Bewertung der Erwartungskonformität.....	143
Abbildung 147. Boxplot-Diagramm für die Bewertung der Sortierung .....	144
Abbildung 148. Boxplot-Diagramm für die Bewertung der Einfachheit .....	145
Abbildung 149. Boxplot-Diagramm für die Bewertung der Innovativität.....	146
Abbildung 150. Boxplot-Diagramm für die Bewertung der Neuartigkeit.....	147
Abbildung 151. Boxplot-Diagramm für die Bewertung der Attraktivität .....	148
Abbildung 152. Boxplot-Diagramm für die Bewertung der Pragmatik.....	149
Abbildung 153. Boxplot-Diagramm für die Bewertung der Sympathie .....	150
Abbildung 154. Boxplot-Diagramm für die Bewertung der Voraussagbarkeit.....	151
Abbildung 155. Boxplot-Diagramm für die Bewertung der Aufgeräumtheit .....	152

Abbildung 156. Boxplot-Diagramm für die Bewertung der Originalität.....	153
Abbildung 157. Boxplot-Diagramm für die Bewertung des Designs .....	154
Abbildung 158. Boxplot-Diagramm für die Bewertung des Werts .....	155
Abbildung 159. Boxplot-Diagramm für die Bewertung der Aufgabenabschlusszeit.....	156

## **Tabellenverzeichnis**

Tabelle 1. Werte der Effektstärke $r$ nach Cohen (1988).....	29
Tabelle 2. Datensätze bei denen der Unterschied statistisch signifikant ist .....	101
Tabelle 3. Datensätze bei denen der Unterschied statistisch nicht signifikant ist.....	102
Tabelle 4. Darstellung der Ergebnisse der deskriptiven Statistik .....	103

# 1. Einleitung

In den letzten Jahren, besonders seit der Corona-Pandemie, hat sich die künstliche Intelligenz stark weiterentwickelt (Büchel et al., 2021), wodurch sich zahlreiche neue Einsatzmöglichkeiten entwickelt haben. KI gilt bereits jetzt als „zukunftsweisende Technologie“<sup>1</sup> (*Was ist künstliche Intelligenz und wie wird sie genutzt?*, 2023) und immer mehr Firmen bringen sie in ihren Arbeitsalltag ein<sup>2</sup> (Berg, 2021; *Künstliche Intelligenz - Perspektive der deutschen Wirtschaft*, 2024). Auch privat nutzen vor allem junge Menschen immer öfter KI-Tools wie ChatGPT, die ein fester Teil des Alltags werden<sup>3</sup> (*Künstliche Intelligenz - Perspektive der deutschen Gesellschaft*, 2024). Diese Entwicklung ist auch im Webdesign zu sehen, wo neue Möglichkeiten entstanden sind. Immer mehr bekannte Website-Baukasten-Anbieter wie IONOS oder WIX bieten KI-Funktionen an, die „im Handumdrehen eine einzigartige, sofort einsatzbereite Website [erstellen]“<sup>4</sup> (*KI Website Erstellen | Kostenlose KI Website in wenigen Minuten*, n.d.). Die KI übernimmt dabei nicht nur die Gestaltung, die Auswahl von Bildern und das Verfassen von Texten, sondern kümmert sich auch um die Suchmaschinenoptimierung<sup>5</sup> (*Homepage-Baukasten » Eigene Webseite mit IONOS*, n.d.).

Diese Entwicklung wirft die Frage auf, ob KI-generierte Applikationen eine günstige und schnelle Alternative zu traditionell manuell gestalteten Applikationen darstellen, die von Designern und Entwicklern erstellt werden müssen. Diese Frage war der Ausgangspunkt meiner Arbeit und brachte mich dazu, eine KI-generierte Applikation mit einer manuell gestalteten Applikation zu vergleichen, insbesondere in Hinblick auf die Benutzererfahrung und -freundlichkeit. Ziel war es herauszufinden, ob KI-generierte Websites tatsächlich eine überzeugende Alternative für Endbenutzer darstellen.

Ich freue mich sehr, die Erkenntnisse dieser Forschung in der vorliegenden Arbeit präsentieren zu können. Die Erstellung dieser Arbeit wäre jedoch ohne die Unterstützung vieler Menschen nicht möglich gewesen, denen ich an dieser Stelle meinen herzlichen Dank aussprechen möchte.

Ein großer Dank geht an Prof. Dr.-Ing. Sebastian Stadler, der mich während des Projekts betreut hat und immer bereit war, meine Fragen zu beantworten. Auch danke ich den Teilnehmern der Usability-Tests für ihre Zeit und hilfreichen Beiträge.

---

<sup>1</sup> <https://www.europarl.europa.eu/topics/de/article/20200827STO85804/was-ist-kunstliche-intelligenz-und-wie-wird-sie-genutzt>

<sup>2</sup> <https://de.statista.com/themen/9400/ki-in-der-deutschen-wirtschaft/#topicOverview>

<sup>3</sup> <https://de.statista.com/themen/9434/ki-in-der-deutschen-gesellschaft/#topicOverview>

<sup>4</sup> <https://de.wix.com/ai-website-builder>

<sup>5</sup> <https://www.ionos.de/websites/homepage-baukasten>

Zu guter Letzt möchte ich meinen Freunden und meiner Familie danken, für ihre emotionale und finanzielle Unterstützung, ohne die diese Arbeit nicht so hätte entstehen können.

## **1.1 Forschungsfrage und Ziel**

In dieser Arbeit wird untersucht, inwiefern sich die Verwendung einer KI-generierten Applikation von einer manuell gestalteten Applikation hinsichtlich der Benutzererfahrung unterscheidet. Um dies zu erforschen, soll vor allem das Verhalten und die Emotionen der Nutzer bei der Interaktion mit einer der beiden Applikationstypen analysiert werden. Zudem wird die Benutzerfreundlichkeit beider Applikationen verglichen, insbesondere in Bezug auf die Effizienz, Effektivität, Zufriedenheit und visuelle Gestaltung.

Das Forschungsthema zielt darauf ab, eine KI-generierte Applikation mit einer manuell gestalteten Applikation zu vergleichen, um herauszufinden, ob und inwieweit KI-generierte Applikationen eine echte Alternative zu manuellen, von Webdesignern entwickelten Applikationen darstellen. Dazu wird zunächst eine Applikation mithilfe einer KI eines gängigen Website-Builders erstellt. Gleichzeitig wird eine manuell gestaltete Applikation entwickelt, die aktuelle Design-Prinzipien, Benutzerfeedback und die Erkenntnisse aus verschiedenen Prototypen berücksichtigt. Ziel ist es, zwei Applikationen zu erhalten, um anschließend einen Usability-Test mit einer begleitenden Umfrage durchführen zu können. Die gesammelten Daten werden nachfolgend analysiert, um beide Ansätze bezüglich der Benutzerfreundlichkeit und Benutzererfahrung vergleichen zu können.

## **1.2 Aufbau der Arbeit**

Die vorliegende Arbeit ist in einen theoretischen und einen praktischen Teil untergliedert. Zunächst werden die theoretischen Grundlagen behandelt, um alle relevanten Begriffe zu definieren. Dazu gehören die Themen User Experience, Usability, User Interface Design, der Human-Centered Design-Prozess und die Gesetze des Webdesigns. Anschließend werden auch Grundlagen der künstlichen Intelligenz behandelt. Dabei wird sowohl auf die Entwicklung und Funktionsweise von KI als auch auf potenzielle Risiken und Gefahren im privaten und beruflichen Umfeld eingegangen. Weiterhin gibt es einen kleinen Ausblick auf die Anwendung von KI im Webdesign. Zuletzt wird die angewendete Datenanalyse beschrieben, mit der die Umfragen der Usability-Tests ausgewertet werden.

Im anschließenden praktischen Teil wird zunächst die Erstellung der manuell gestalteten Applikation vorgestellt, die mithilfe der Methoden des Human-Centered Design-Prozesses entwickelt wurde. Dieser ist in der vorliegenden Arbeit in vier Phasen unterteilt. Die erste Phase, dem Verstehen und Beschreiben des Nutzungskontextes, dient dem Verständnis der Bedürfnisse, Wünsche und Herausforderungen der Nutzer. Dies bildet die Grundlage für spätere Designentscheidungen und der Entwicklung einer verbesserten sowie nutzerzentrierten Applikation. In der darauffolgenden Phase werden die Nutzungsanforderungen spezifiziert, indem die gewonnenen Erkenntnisse aus der vorherigen Phase definiert und festgehalten werden. Zudem werden klare Problemstellungen formuliert, die als Basis für die weiteren Entwicklungsschritte dienen. Es folgt die kreative Phase der Lösungsentwicklung, in der Ideen und Lösungsansätze zu den zuvor definierten Problemen erarbeitet werden. Anschließend werden die Ideen geordnet und ein erster grober Aufbau der Website entsteht. Danach werden wichtige Design-

Entscheidungen getroffen, beispielsweise die Wahl des Entwicklungsprogramms sowie die Gestaltung der Schrift, Farben, Buttons und Icons. Zuletzt werden erste ausgearbeitete Versionen der bisherigen Ideen erstellt. Diese Prototypen ermöglichen frühe Nutzertests und ein wertvolles Feedback für weitere Verbesserungen.

Neben der manuellen Gestaltung wird im praktischen Teil auch die Erstellung der KI-generierten Applikation erläutert. Hierbei wird das verwendete Programm WIX vorgestellt und der Entwicklungsprozess der Applikation dokumentiert. Dabei wurde die KI-gestützte Erstellung durch eine Person ohne Vorkenntnisse im Webdesign durchgeführt.

Zum Abschluss werden beide Applikationen im Rahmen eines moderierten Usability-Tests getestet und durch Umfragen bewertet. Die gesammelten Daten werden anschließend ausgewertet, um einen fundierten Vergleich zwischen den beiden Ansätzen zu ermöglichen.

## 2. Theorie

### 2.1 Grundlagen Webdesign

In den folgenden Unterkapiteln werden grundlegende Begriffe und Gesetze aus dem Bereich Webdesign genauer behandelt, um einen guten Überblick für die Anforderungen an eine nutzerfreundliche Website zu erhalten.

#### 2.1.1 User Experience

Der Begriff User Experience, kurz UX, kann als „Benutzererlebnis“ oder „Benutzererfahrung“ ins Deutsche übersetzt werden und wird in der DIN-EN-ISO-9241-210 definiert. Dabei bezieht sich die UX auf „die Wahrnehmungen und die Reaktionen eines Benutzers, die sich aus der Nutzung und/oder der erwarteten Nutzung eines interaktiven Systems ergeben“ (Geis & Tesch, 2019, S. 17). Während die nachfolgend beschriebene Usability vor allem die Effektivität, Effizienz und den Spaßfaktor der Benutzer während einer Interaktion misst, betrachtet die User Experience das gesamte Nutzererlebnis vor, während und nach der Nutzung (Jacobsen & Meyer, 2022). Dabei spielen auch Emotionen, Vorstellungen, Vorlieben, Wahrnehmungen, physiologische und psychologische Reaktionen, Verhaltensweisen und Leistungen eine Rolle (Quirmbach, 2012, S. 58).



Abbildung 1. Unterschied zwischen User Experience und Usability (Jacobsen & Meyer, 2022, S. 33)

Wie in Abbildung 2 ersichtlich wird, beginnt die User Experience schon vor der Nutzung eines Produkts und hängt stark von den Erwartungen der Nutzer ab. Ob diese Erwartungen erfüllt oder enttäuscht werden, hat ebenso Auswirkungen auf das Benutzererlebnis wie die Nutzung selbst. Ein positives Markenimage kann beispielsweise dabei helfen schon im Vorfeld eine positive Erwartungshaltung zu erzeugen und die User Experience entsprechend zu verbessern. Während der Nutzung sind Faktoren wie eine einfache und schnelle Bedienbarkeit, die Entdeckung unerwarteter hilfreicher Funktionen sowie ein ansprechendes Design entscheidend. Nach der Nutzung ist es bedeutend, ob sich die Nutzer mit dem Produkt identifizieren können, dadurch eine emotionale Bindung aufzubauen und somit ein gutes Benutzererlebnis haben. Negative Erfahrungen, wie längere Lieferzeiten, können das Erlebnis später beeinträchtigen. Im Gegensatz dazu beschäftigt sich die Usability nur mit der tatsächlichen Nutzung der Anwendung (Jacobsen & Meyer, 2022).

Laut Quirmbach (2012) ist die User Experience anfangs neutral und lässt keine direkten Rückschlüsse auf die Qualität des Erlebnisses zu. Erst durch den Einsatz verschiedener Bewertungsmethoden kann bewertet werden, ob das Erlebnis positiv oder negativ sowie erfolgreich oder nicht erfolgreich war. Man kann dabei objektive Kriterien wie die Attraktivität des Produkts, die Einfachheit der Bedienung und die Emotionen, die bei der Nutzung entstehen, betrachten. Auch wie neu oder innovativ ein Produkt ist, sowie der Spaßfaktor spielen eine Rolle. Ein weiterer wichtiger Aspekt ist die Usability, welche die Effektivität, Effizienz und Zufriedenheit der Benutzer bewertet. Zusätzlich wird die Utility der Funktionalitäten berücksichtigt, also der Nutzen und die Zweckmäßigkeit der angebotenen Funktionen (Quirmbach, 2012, S. 58). Diese Faktoren sind wichtig, um die Qualität der Benutzererfahrung gut bewerten und gegebenenfalls verbessern zu können.

## 2.1.2 Usability

Der Begriff Usability kann mit „Verwendbarkeit“, „Gebrauchstauglichkeit“ oder „Benutzbarkeit“ übersetzt werden. Laut der DIN-EN-ISO-9241-11 beschreibt die Usability, wie ein Produkt von bestimmten Nutzern in einem bestimmten Nutzungskontext effektiv, effizient und zufriedenstellend genutzt werden kann, um ein festgelegtes Ziel zu erreichen. Man kann die Usability nach verschiedenen Aspekten bewerten, wie dem ersten Eindruck, der Offensichtlichkeit und Eindeutigkeit des Verwendungszwecks eines Elements, die intuitive Bedienbarkeit, die klar erkennbaren physischen und logischen Eigenschaften der Bedienelemente sowie die einfache und langfristige Erlernbarkeit unbekannter Funktionen. Weitere Kriterien umfassen das Tempo der Aufgabenerfüllung, die Erfüllung der Funktionserwartung eines Elements, die Einhaltung von Reaktions- und Konvergenzerwartungen sowie die subjektive Benutzerzufriedenheit. Insgesamt misst die Usability somit die Qualität des zuvor beschriebenen User Interface (Chlebek, 2011).

Jakob Nielsens (1994) zehn Heuristiken für interaktives Design legen den Fokus auf die Kommunikation zwischen dem System und den Nutzern. Der Status des Systems sollte dabei immer sichtbar sein und klare Rückmeldungen geben, um den Benutzer stets über den aktuellen Stand des Systems zu informieren. Dies umfasst unter anderem Statusanzeigen, Fortschrittsindikatoren und klare sowie rechtzeitige Reaktionen des Systems auf Eingaben. Das System sollte so gestaltet sein, dass es der realen Welt ähnelt, die Sprache der Nutzer spricht und bekannte Begriffe verwendet, um eine intuitive Bedienung zu ermöglichen. Zudem sollte es an bereits bekannten mentalen Modellen anknüpfen und bestehendes Nutzerwissen verwenden. Nutzer sollten die Kontrolle über ihre Aktionen haben, mit Funktionen wie „rückgängig“ und „wiederholen“ sowie klaren Ausgängen und der Möglichkeit, Aktionen abzubrechen oder neu zu ordnen. Konsistenz und Standards erfordern, dass dieselben Dinge auf gleiche Weise ausgedrückt und visuell einheitlich dargestellt werden, um eine benutzerfreundliche Erfahrung zu gewährleisten. Weiterhin sollten Fehler schon im Vorfeld vermieden werden, indem Systeme so gestaltet sind, dass sie potenzielle Fehlerquellen erkennen und die Sprache der Benutzer verstehen. Auch Hilfen und Hinweise sind wichtig, um die Nutzer zu unterstützen. Die Heuristik „Erkennung statt Erinnerung“ fördert die Usability, indem Aktionen und Optionen sichtbar gemacht werden, wodurch die Gedächtnislast für die Nutzer minimiert wird. Dies umfasst auch die Bereitstellung von Listen zur Auswahl und bereits bekannte Funktionen aus anderen Apps. Flexibilität und Effizienz erfordern, dass Nutzer durch Tastenkürzel und anpassbare Funktionen schneller arbeiten können, um damit die Usability zu optimieren. Die letzten drei Faktoren betonen die Wichtigkeit eines ästhetischen und minimalistischen Designs, klar strukturierte Funktionen, die leicht zu unterscheiden sind und die Verwendung von Standardwerten, damit Nutzer keine Informationen mehrfach eingeben müssen.



Abbildung 2. Nielsen's 10 Usability Heuristiken (Nielsen, 1994), Grafik: Agente<sup>6</sup> (2019)

### 2.1.3 User Interface

Der Begriff User Interface, oder kurz UI, bedeutet auf Deutsch so viel wie „Benutzeroberfläche“, „Benutzerschnittstelle“ oder „Mensch-Maschine-Schnittstelle“. In der DIN-EN-ISO-9241-110 wird das User Interface definiert als „Alle Bestandteile eines interaktiven Systems (Software oder Hardware), die Informationen und Steuerelemente zur Verfügung stellen, die für den Benutzer notwendig sind, um eine bestimmte Arbeitsaufgabe mit dem interaktiven System zu erledigen.“ (Chlebek, 2011, S. 15). Das heißt, es geht um die Bedienelemente einer Anwendung, ihre Anordnung und wie sie miteinander interagieren. Bei Applikationen stellen diese Bedienelemente sämtliche Eingabe- und Ausgabeformen dar, wie zum Beispiel Eingabefelder, Anzeigen, Töne und haptische Signale (Chlebek, 2011). Ein Beispiel für haptische Signale könnten Vibrationen sein, die ein Smartphone bei einer eingegangenen Nachricht am Handy erzeugt. Bei der Gestaltung des User Interface, muss man also alle Handlungsmöglichkeiten und deren Feedback berücksichtigen, die ein Nutzer braucht, um seine Aufgaben zu erledigen (Geis & Tesch, 2019).

Der Begriff User Interface bezieht sich demnach nicht ausschließlich auf das Aussehen einer Anwendung. Thesmann (2009) betont, dass die Aufgabe von UI darin besteht, „dem Anwender die für die

---

<sup>6</sup> <https://www.epicpixels.com/items/set-of-free-posters-for-user-interface-design>

Bewältigung einer Aufgabe erforderliche Funktionalität eines Werkzeugs auf ergonomische Weise zugänglich zu machen“ (Thesmann, 2009, S. 7).

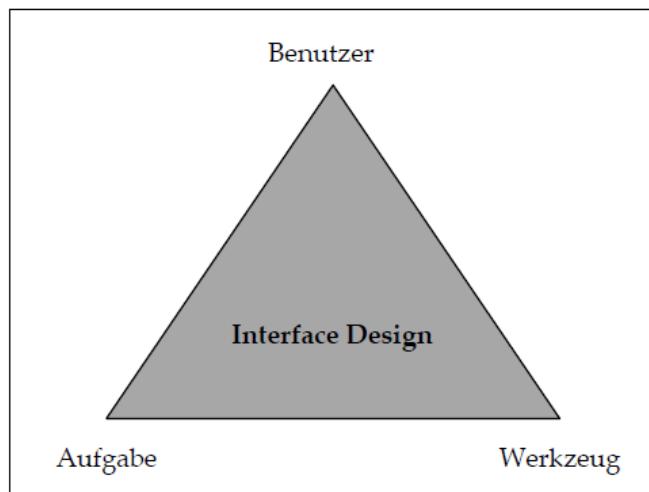


Abbildung 3. Ontologisches Designdiagramm von Gui Bonsiepe von 1996 (Thesmann, 2009, S. 7)

In diesem Zusammenhang greift Thesmann (2009) auch das Ontologische Designdiagramm von Gui Bonsiepe aus seinem Buch *Interface: Design neu begreifen* (1996) auf. Dieses verdeutlicht, dass das User Interface Design als zentrales Element fungiert, das den Benutzer, die zu lösende Aufgabe und das dafür benötigte Werkzeug miteinander verbindet. Um eine optimale Nutzung zu ermöglichen, müsse das Design auf die Anforderungen der Aufgabe sowie auf die Fähigkeiten, Erfahrungen und Präferenzen des Benutzers und die Funktionalität des Werkzeugs abgestimmt sein (Thesmann, 2009).

#### 2.1.4 Gesetze für Webdesign

Im Bereich Webdesign gibt es einige Gesetze und Prinzipien zu beachten, um ein überschaubares und einfach verwendbares Design zu garantieren. Im Folgenden werden einige Gesetze vorgestellt.

##### Das Gesetz der Nähe

Das Gesetz der Nähe ist eines der Gestaltgesetze aus der Wahrnehmungspsychologie. Es besagt, dass unsere visuelle Wahrnehmung Objekte, die sich räumlich nahe beieinander befinden, automatisch als zusammengehörig gruppiert. Der Raum zwischen Elementen, auch Weißraum genannt, unterstützt dabei die Ordnung und erleichtert die Orientierung. Zusammengehörende Elemente sollten demnach nah zusammen platziert werden. Umgekehrt empfiehlt es sich, nicht zusammengehörige Bereiche, wie die Navigation, den Hauptinhalt und die Fußzeile, durch ausreichend Weißraum deutlich voneinander abzugrenzen (Hahn, 2020).

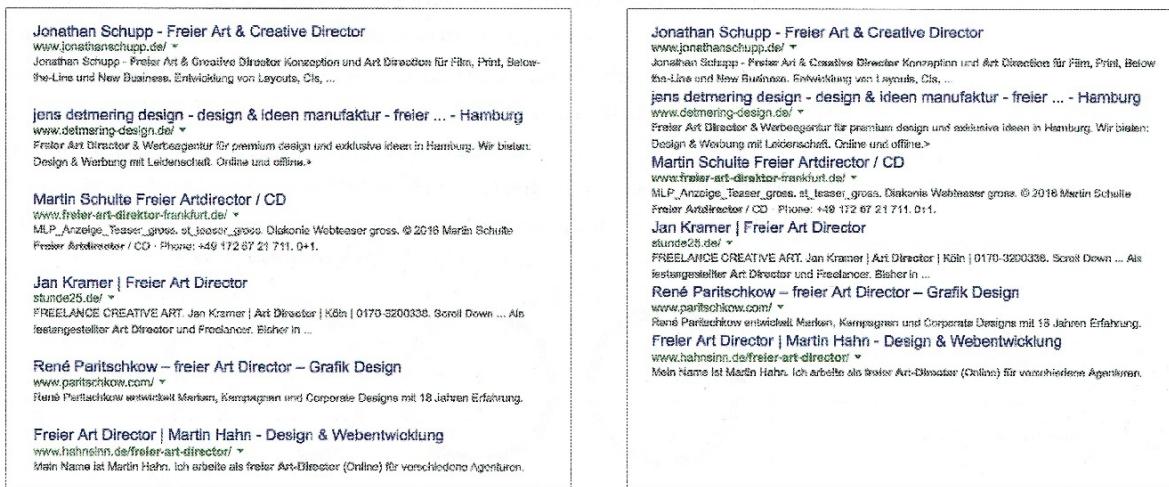


Abbildung 4. Gesetz der Nähe am Beispiel des Google-Suchergebnis (Hahn, 2020, S. 223)

In Abbildung 4 ist die Wirkung von Weißraum deutlich zu erkennen. Auf der rechten Seite ist das originale Ergebnis einer Google-Suche dargestellt, bei dem die einzelnen Treffer durch den gezielten Einsatz von Weißraum voneinander abgegrenzt sind. Dies ermöglicht eine klare visuelle Trennung der Informationen, wodurch sie für die Nutzer leicht erfassbar und übersichtlich angeordnet sind. Im Gegensatz dazu zeigt die linke Darstellung eine Version ohne Weißraum, in der die Treffer kaum voneinander abgegrenzt sind und somit schwerer voneinander zu unterscheiden sind (Hahn, 2020).

### **Das Gesetz der Geschlossenheit**

Das Gesetz der Geschlossenheit ist ein Prinzip der Gestaltpsychologie und beschreibt die Tendenz unseres Gehirns, unvollständige Formen oder Strukturen als vollständig wahrzunehmen. Wenn einzelne Elemente eine erkennbare Form andeuten, neigt unsere Wahrnehmung dazu, die fehlenden Informationen automatisch zu ergänzen. Dadurch entstehen in unserem Kopf „geschlossene“ Figuren, selbst wenn diese physisch gar nicht vollständig dargestellt sind.

Dieses Gesetz wird im Webdesign gezielt eingesetzt, um Einheiten zu schaffen, Informationen zu strukturieren und Zusammenhänge deutlich zu vermitteln. Dies kann zum einen durch das zuvor beschriebene Gesetz der Nähe erfolgen oder durch das Einrahmen von Elementen mit Linien oder Kästen, durch farblich hinterlegte Bereiche oder durch eine einheitliche Gestaltung. Das Gesetz der Geschlossenheit ergänzt somit ideal das Gesetz der Nähe (Hahn, 2020).



Abbildung 5. Gesetz der Geschlossenheit am Beispiel einer Anmeldeseite (Hahn, 2020, S. 225)

In Abbildung 5 wird zum einen das Gesetz der Nähe durch die gezielten Abstände zwischen den Elementen deutlich. Zum anderen lässt sich auch das Gesetz der Geschlossenheit erkennen, da alle drei Blöcke dem gleichen Aufbau aus Bild, Überschrift und Text folgen. Dadurch werden sie als abgeschlossene Einheiten wahrgenommen, was eine klare visuelle Struktur erzeugt. Dies unterstützt die Übersichtlichkeit und erleichtert die Informationsverarbeitung für die Nutzer (Hahn, 2020).



Abbildung 6. Nicht berücksichtigtes Gesetz der Geschlossenheit (Hahn, 2020, S. 226)

In Abbildung 6 wird das Gesetz der Geschlossenheit nicht berücksichtigt, da der Aufbau der einzelnen Blöcke variiert. Durch die nicht einheitliche Gestaltung sind zusammengehörige Elemente nur schwer als solche erkennbar, was die visuelle Struktur auflöst und die Informationsverarbeitung der Nutzer erheblich erschwert (Hahn, 2020).

## Gesetz der Ähnlichkeit

Das Gesetz der Ähnlichkeit ist ein Gestaltprinzip aus der Wahrnehmungspsychologie. Es besagt, dass Elemente, die sich in bestimmten Merkmalen ähneln, etwa in Farbe, Material beziehungsweise Textur, Helligkeit, Form, Größe, Position, Bewegungsrichtung oder Geschwindigkeit, von unserer Wahrnehmung automatisch als zusammengehörig gedeutet werden. Diese visuelle Gruppierung erfolgt intuitiv, selbst wenn die Elemente räumlich voneinander getrennt sind. „Je mehr Gemeinsamkeiten zwei Elemente aufweisen, desto stärker ist die Wahrnehmung als Einheit“ (Hahn, 2020, S. 227).

Im Kontext des Webdesign spielt dieses Prinzip eine zentrale Rolle. Es ermöglicht, gleichartige Elemente wie Buttons oder Überschriften optisch als funktional zusammengehörig zu kennzeichnen. So wird deutlich, dass sie dieselbe Bedeutung oder Funktion besitzen. Gleichzeitig unterstützt das Gesetz der Ähnlichkeit die Wiedererkennbarkeit über verschiedene Unterseiten hinweg und verbessert damit die Benutzerfreundlichkeit. Wichtig ist dabei, eine Balance zwischen gestalterischer Konsistenz und visueller Abwechslung zu wahren, um ein monotoner sowie vorhersehbares Design und die damit einhergehende Langeweile zu vermeiden (Hahn, 2020).

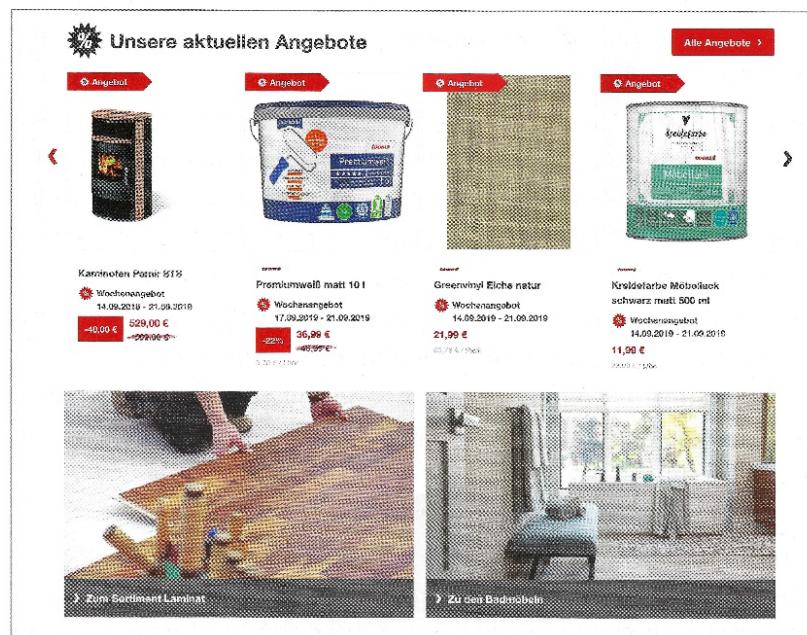


Abbildung 7. Gesetz der Ähnlichkeit am Beispiel von toom.de (Hahn, 2020, S. 227)

In Abbildung 7 werden mehrere Elemente trotz des großen Abstands zueinander als zusammengehörig wahrgenommen, da sie visuelle Gemeinsamkeiten aufweisen. So haben beispielsweise die Angebotsbanner bei allen Produkten die gleiche Größe, Farbe und Position, wodurch eine visuelle Gruppierung entsteht und ihre gemeinsame Funktion erkennbar wird (Hahn, 2020).

## Gesetz der Erfahrung

Das Gesetz der Erfahrung ist ein Gestaltgesetz aus der Wahrnehmungspsychologie. Es besagt, dass unsere Wahrnehmung stark von unseren früheren Erfahrungen und Begegnungen beeinflusst wird. Wenn wir bestimmte Formen, Strukturen oder Anordnungen bereits kennen, erkennen wir sie auch dann wieder, wenn sie unvollständig oder abstrahiert dargestellt sind.

Im Designkontext bedeutet das, dass bekannte Platzierungen und Gestaltungselemente gezielt genutzt werden können, um Orientierung und Verständnis zu fördern. So erwarten Nutzer beispielsweise das Logo einer Website in der oberen linken Ecke und die Navigation direkt darunter oder oben rechts. Auch das Diskettensymbol zum Speichern von Daten ist ein typisches Beispiel für ein vertrautes Element, das unabhängig von seiner tatsächlichen Funktion intuitiv verstanden wird. In Abbildung 8 lassen sich weitere Symbole erkennen, deren Bedeutung aufgrund ihrer Bekanntheit unmittelbar erkannt wird, was die Interaktion für den Nutzer erleichtert (Hahn, 2020).

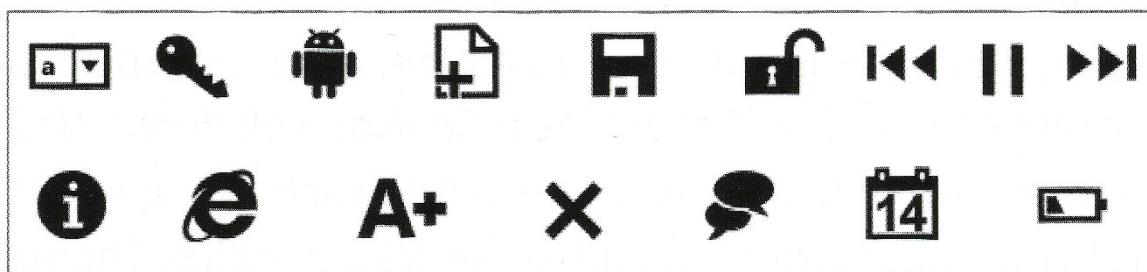


Abbildung 8. Gesetz der Erfahrung anhand von Symbolen (Hahn, 2020, S. 229)

Nutzer interpretieren neue Informationen also oft basierend auf dem, was sie bereits kennen. Je vertrauter ein Element wirkt, desto leichter ist es für sie zu verstehen und zu bedienen. Eine Abweichung von diesem Gesetz sollte im begrenzten Rahmen erfolgen, da dies zwar als Innovation und Abgrenzung von anderen Websites gelten kann, jedoch auch die Gefahr erzeugt, den Nutzer zu frustrieren, was die Usability und User Experience stark ins Negative beeinflussen kann (Hahn, 2020).

## Jakobs Gesetz

Ähnlich wie das Gesetz der Erfahrung funktioniert auch Jakobs Gesetz. Es wurde von Jakob Nielsen im Jahr 2000 entwickelt, einem Experten im Bereich der Usability. Es besagt, dass Nutzer die meiste Zeit auf anderen Websites verbringen und daher erwarten, dass sich neue Websites ähnlich verhalten wie die, die sie bereits kennen. Nutzer bevorzugen Vertrautheit, weil sie bereits eine Vorstellung davon haben, wie etwas zu nutzen ist. Das heißt also je weniger Anstrengung Nutzer aufbringen müssen, um eine Benutzeroberfläche zu verstehen, desto mehr können sie sich auf ihre eigentliche Aufgabe konzentrieren. Wenn ihnen der Weg zum Ziel so einfach wie möglich gemacht wurde, steigt die Wahrscheinlichkeit, dass sie erfolgreich und zufrieden sind (Yablonski, 2024).

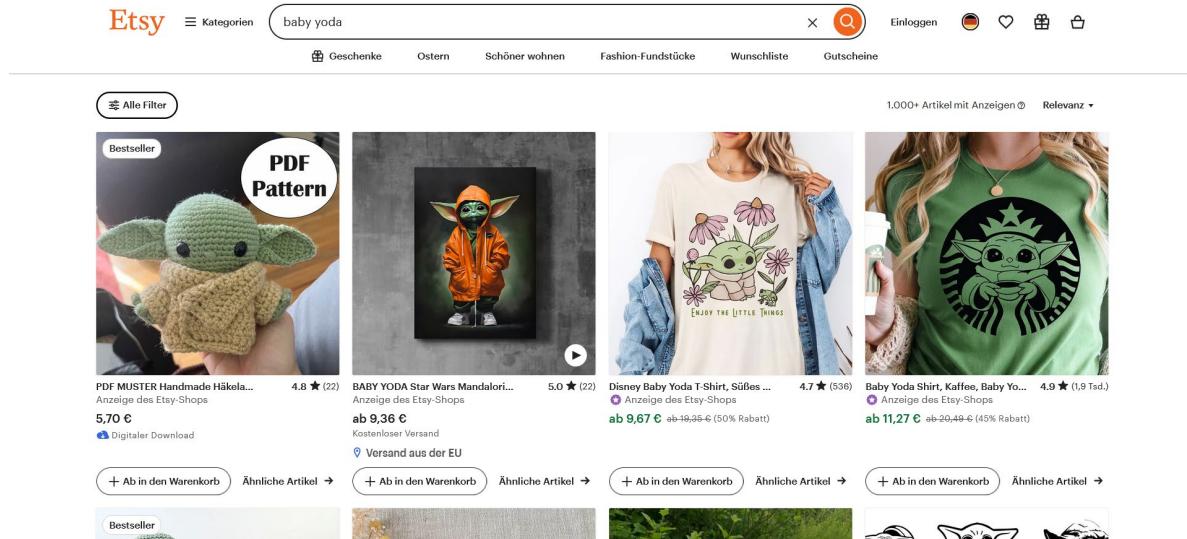


Abbildung 9. Jakobs Law anhand von Etsy erklärt<sup>7</sup> (Quelle: Etsy, 2025)

In Abbildung 9 ist die Verkaufsseite von Etsy dargestellt, die gezielt vertraute mentale Modelle nutzt, um die Benutzeroberfläche möglichst intuitiv und benutzerfreundlich zu gestalten. Ziel ist es, den Nutzern die Interaktion zu erleichtern, sodass sie sich auf den Kaufprozess konzentrieren können, ohne neue Bedienmuster erlernen zu müssen. Der Aufbau der Produktanzeige orientiert sich dabei an anderen Websites. Zunächst wird das Produkt über ein Bild präsentiert, dann folgt der Produktnname, der Preis, der Versand und anschließend Aktionen wie „In den Warenkorb“ oder „Ähnliche Produkte“ (Yablonski, 2024).

### Gesetz der Kontinuität

Das Gesetz der Kontinuität ist ein Gestaltprinzip der Wahrnehmungspsychologie. Es besagt, dass wir Elemente, die entlang einer Linie oder Kurve angeordnet sind, als zusammengehörig wahrnehmen. Unser Gehirn neigt dazu Linien als eine Einheit zu interpretieren, auch wenn sie unterbrochen werden.

Im Kontext des Webdesign bedeutet das, dass der Nutzer Sinnzusammenhänge erstellt und Inhalte selbstständig ergänzt (Hahn, 2020).



Abbildung 10. Gesetz der Kontinuität am Beispiel von Seitenzahlen (Hahn, 2020, S.).

<sup>7</sup> [https://www.etsy.com/de-en/search?q=baby%20yoda&ref=search\\_bar](https://www.etsy.com/de-en/search?q=baby%20yoda&ref=search_bar)

Anhand von Abbildung 10 lässt sich das Gesetz der Kontinuität veranschaulichen. Dargestellt sind die ersten und letzten Seitenzahlen einer Website, wobei die dazwischenliegenden Zahlen nicht angezeigt werden. Die Nutzer ergänzen diese fehlenden Informationen jedoch automatisch, da sie eine fortlaufende Reihenfolge erwarten und diese selbstständig fortsetzen.

### Fitts' Gesetz

Fitts' Gesetz, entwickelt von dem amerikanischen Psychologen Paul Fitts im Jahr 1945, beschreibt, wie schnell ein Nutzer ein Ziel mit einem Zeigegerät, etwa einem Mauszeiger oder dem Finger, erreichen kann. Es „besagt, dass die Zeit, die ein Benutzer braucht, um mit einem Objekt zu interagieren, im Verhältnis zu dessen Größe und Entfernung steht“ (Yablonski, 2024, S. 31). Je größer ein interaktives Element ist und je näher es sich zur aktuellen Position des Zeigers oder Fingers befindet, desto schneller und einfacher kann es ausgewählt werden. Umgekehrt verlängert sich die benötigte Zeit, wenn das Ziel kleiner oder weiter entfernt ist. Im Kontext der Gestaltung von Benutzeroberflächen bedeutet dies, dass interaktive Elemente wie Buttons oder Eingabefelder groß genug und gut erreichbar platziert sein sollten (Yablonski, 2024).

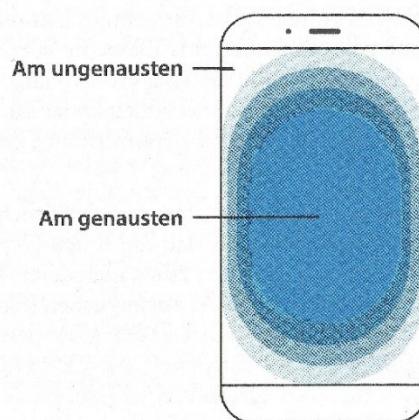


Abbildung 11. Touch-Genauigkeit bei Smartphones (Yablonski, 2024, S. 36)

In Abbildung 11 lässt sich klar erkennen, an welchen Stellen das Smartphone am genauesten auf Berührungen reagiert und wo die Elemente optimal platziert werden sollten, um für den Nutzer bestmöglich erreichbar zu sein. Dies verbessert die Interaktion mit der Anwendung (Yablonski, 2024).

Address 1 265 Main St	Address 2 Apt 102
City Portland	Zip Code 97201
<input type="button" value="Submit"/>	

Abbildung 12. Fitts' Gesetz anhand eines Formularfeldes (Yablonski, 2024, S. 37)

Ein gutes Beispiel für Fitts' Gesetz ist in Abbildung 12 zu sehen. Der Senden-Button ist nicht weit von dem letzten Eingabefeld entfernt platziert. Dadurch entsteht einerseits eine visuelle Verbindung

zwischen den beiden Aktionen und andererseits wird sichergestellt, dass der Nutzer nur eine kurze Strecke zurücklegen muss, um das Ziel abzuschließen (Yablonski, 2024).

### **Millers Gesetz**

Millers Gesetz geht auf den Kognitionspsychologen George Miller zurück. In seiner Arbeit "*The Magical Number Seven, Plus or Minus Two*" von 1956 stellte er fest, dass das Kurzzeitgedächtnis des Menschen im Durchschnitt nur etwa sieben Informationseinheiten gleichzeitig verarbeiten und speichern kann, mit einer Schwankungsbreite von etwa zwei Einheiten. Das bedeutet, dass Menschen typischerweise nur fünf oder neun Elemente gleichzeitig im Kurzzeitgedächtnis behalten können. Im Kontext des Webdesigns bedeutet dies jedoch nicht, sich von diesem Gesetz zu sehr einschränken zu lassen. Es soll lediglich dafür gesorgt werden, dass Informationsmengen reduziert, strukturiert und sinnvoll gruppiert werden sollten, um die kognitive Belastung für die Nutzer möglichst gering zu halten. Dies sollte durch Gestaltungselemente wie Farbe, Skalierung, Unterteilung und Abstände erfolgen (Yablonski, 2024).

A wall of text is an excessively long post to a noticeboard or talk page discussion, which can often be so long that some don't read it. Some walls of text are intentionally disruptive, such as when an editor attempts to overwhelm a discussion with a mass of irrelevant kilobytes. Other walls are due to lack of awareness of good practices, such as when an editor tries to cram every one of their cogent points into a single comprehensive response that is roughly the length of a short novel. Not all long posts are walls of text; some can be nuanced and thoughtful. Just remember: the longer it is, the less of it people will read. The chunk-o'text defense (COTD) is an alleged wikilawyering strategy whereby an editor accused of wrongdoing defends their actions with a giant chunk of text that contains so many diffs, assertions, examples, and allegations as to be virtually unanswerable. However, an equal-but-opposite questionable strategy is dismissal of legitimate evidence and valid rationales with a claim of "text-walling" or "TL;DR". Not every matter can be addressed with a one-

Abbildung 13. Millers Gesetz: Beispiel einer Textwüste<sup>8</sup> (*Chapter 4. Miller's Law*, n.d.)

In Abbildung 13 ist eine sogenannte „Textwüste“ zu sehen, die weder eine klare Hierarchie noch eine sinnvolle Formatierung aufweist und zudem keine angemessene Zeilenlänge besitzt. Dies erschwert das Erfassen und Verarbeiten des Inhalts, was zu einer höheren kognitiven Belastung des Nutzers führt und sich negativ auf die User Experience auswirkt (Yablonski, 2024).

---

<sup>8</sup> <https://www.oreilly.com/library/view/laws-of-ux/9781492055303/ch04.html>

## Wall of Text

A wall of text is an excessively long post to a noticeboard or talk page discussion, which can often be so long that some don't read it.

### Types

Some walls of text are intentionally disruptive, such as when an editor attempts to overwhelm a discussion with a mass of irrelevant kilobytes. Other walls are due to lack of awareness of good practices, such as when an editor tries to cram every one of their cogent points into a single comprehensive response that is roughly the length of a short novel. Not all long posts are walls of text; some can be nuanced and thoughtful.

Just remember: **the longer it is, the less of it people will read.**

Abbildung 14. Millers Gesetz: Verbesserte Textwüste<sup>9</sup> (*Chapter 4. Miller's Law*, n.d.)

Im Gegensatz zur vorherigen Abbildung zeigt Abbildung 14 einen formatierten Text mit klaren Hierarchien und einer angemessenen Zeilenlänge. Durch den Einsatz von Überschriften, Unterüberschriften, Leeräumen zwischen den Textabschnitten und einer verkürzten Textlänge wurde die Lesbarkeit deutlich verbessert, wodurch der Nutzer den Inhalt wesentlich leichter erfassen und verarbeiten kann (Yablonski, 2024).

## Hicks Gesetz

Hicks Gesetz, entwickelt von den Psychologen William Edmund Hick und Ray Hyman im Jahr 1952, beschreibt den Zusammenhang zwischen der Anzahl an Wahlmöglichkeiten und der Zeit, die ein Mensch benötigt, um eine Entscheidung zu treffen. Je mehr Optionen zur Auswahl stehen, desto länger dauert die Entscheidungsfindung. Im Kontext des Webdesigns bedeutet dies, dass „die Zeit, die der Nutzer für die Interaktion mit einer Schnittstelle benötigt, [...] direkt von der Anzahl der ihm zur Verfügung stehenden Auswahlmöglichkeit [abhängt]“ (Yablonski, 2024, S. 54). Eine überladene Benutzeroberfläche kann somit zu längeren Verarbeitungszeiten führen und erschwert es Nutzern relevante Inhalte oder Funktionen zu identifizieren. Yablonski (2024) betont „Je weniger er [der Nutzer] darüber nachdenken muss, auf welche Weise er sein Ziel erreichen kann, desto wahrscheinlicher ist es, dass er es erreichen wird“ (S. 63). Zur Erstellung einer übersichtlichen sowie nutzerfreundlichen Informationsstruktur können Methoden wie das Card Sorting eingesetzt werden, bei denen Inhalte aus Nutzersicht in logische Gruppen eingeteilt werden. Zusätzlich kann das Prinzip der Vergrößerung hilfreich sein. Dabei wird die Benutzeroberfläche, vor allem Inhalte wie Icons, so weit wie möglich vereinfacht, ohne die

<sup>9</sup> <https://www.oreilly.com/library/view/laws-of-ux/9781492055303/ch04.html>

Verständlichkeit negativ zu beeinflussen und dabei eine erhöhte kognitive Belastung hervorzurufen (Yablonski, 2024).

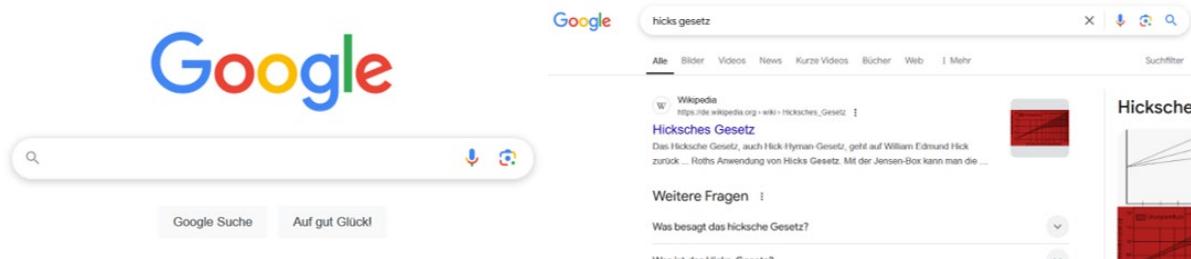


Abbildung 15. Hicks Gesetz anhand der Google-Suche<sup>10</sup> (Quelle: Google, 2025)

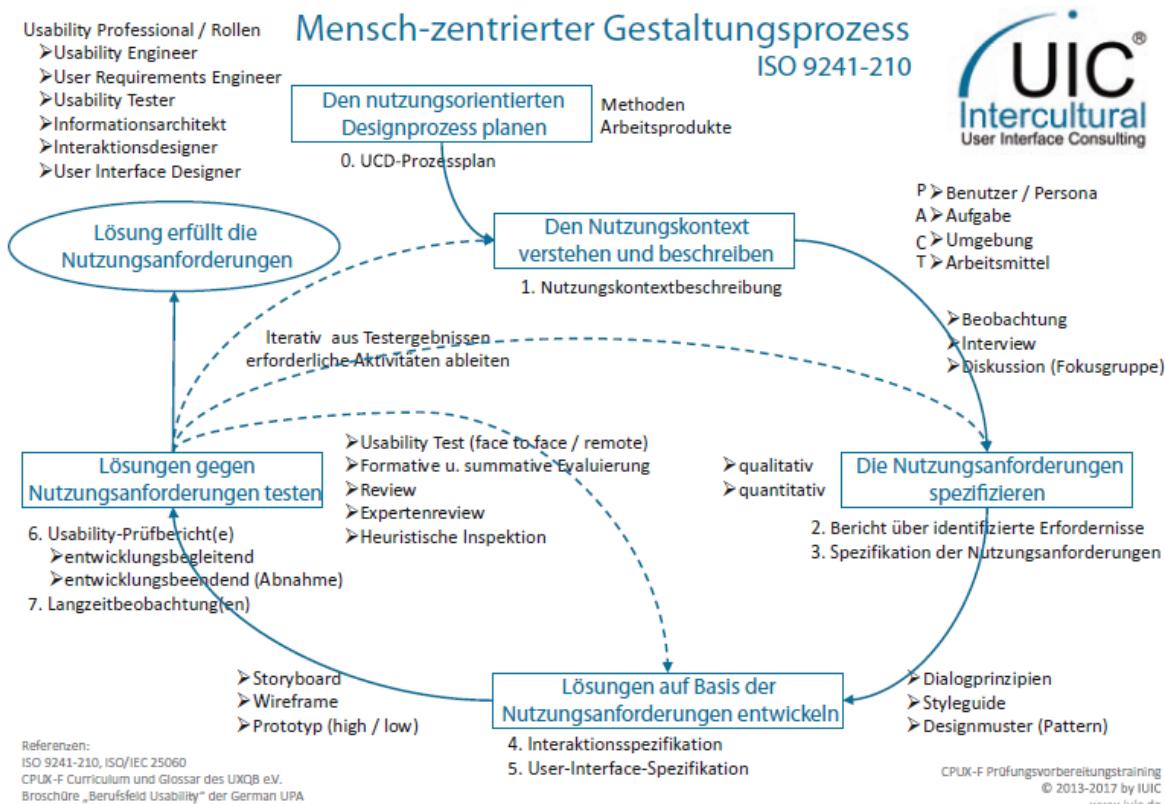
In Abbildung 15 ist ein gutes Beispiel für Hicks Gesetz zu erkennen. Zunächst werden dem Nutzer, wie in dem linken Bild zu erkennen ist, nur eine begrenzte Anzahl an Auswahlmöglichkeiten angeboten, um es dem Nutzer zu erleichtern sich auf das aktuelle Ziel zu konzentrieren und nicht durch eine Vielzahl von Optionen überfordert zu werden. Im rechten Bild ist das Suchergebnis sowie die Möglichkeit, diese durch Filter weiter einzuschränken, dargestellt. Indem dem Nutzer zur richtigen Zeit die passenden Auswahlmöglichkeiten präsentiert wurden, wurde die User Experience verbessert.

---

<sup>10</sup> [https://www.google.com/?zx=1746100392879&no\\_sw\\_cr=1](https://www.google.com/?zx=1746100392879&no_sw_cr=1)

## 2.1.5 Der Human-Centered Design-Prozess

Der Human-Centered Design-Prozess, ins Deutsche übersetzt als „menschzentrierter Gestaltungsprozess“, basiert auf den Anforderungen und Empfehlungen aus dem Teil 210 der DIN-EN-ISO-9241 festgelegt sind. Der Prozess besteht aus vier Phasen, die nötig sind, um die Benutzerfreundlichkeit interaktiver Systeme während der Entwicklung zu gewährleisten.



**Abb. 4.6** Mensch-Zentrierter Gestaltungsprozess in Anlehnung an ISO 9241-210. (Quelle: IUC)

Abbildung 16. Phasen des Human-Centered Design-Prozess (Heimgärtner, 2017, S. 106)

In Abbildung 16 wird deutlich, dass der Human-Centered Design-Prozess aus vier zentralen Phasen besteht. Dem Verstehen und Beschreiben des Nutzungskontextes, der Spezifizierung der Nutzungsanforderungen, der Entwicklung von Lösungen auf Basis dieser Anforderungen sowie dem Testen der Lösungen. Darüber hinaus zeigt die Abbildung jedoch, dass es zusätzlich zu den vier Phasen noch zwei weitere gibt. Burghardt et al. beschreiben diese als einen zusätzlichen „Start- (Feststellen der Notwendigkeit einer benutzerzentrierten Gestaltung) und einen Endpunkt (Das System erfüllt die festgelegten Anforderungen an Funktion, Organisation und Benutzerbelangen)“ (Burghardt et al., 2011, S. 365). Zu Beginn des Designprozesses wird zunächst geplant und überprüft, ob sich eine spezifische Aufgabe überhaupt mithilfe eines benutzerzentrierten Ansatzes lösen lässt. Am Ende des Prozesses erfolgt eine langfristige Evaluation, bei der geprüft wird, ob das endgültige System die festgelegten Anforderungen an Funktion, Organisation und Benutzerbelange im realen Anwendungskontext erfüllt. Langzeitbeobachtungen ermöglichen dabei eine umfassendere Bewertung als im Prototypenstadium (Burghardt et

al., 2011). In dieser Arbeit entfällt jedoch dieser letzte Schritt, da der Prototypenstadium nicht verlassen wird.

## 1. Nutzungskontext verstehen und beschreiben

In der ersten Phase wird der Nutzungskontext analysiert und definiert. Dabei werden Benutzermerkmale, Arbeitsaufgaben sowie die organisatorische und physische Umgebung berücksichtigt. Diese Informationen werden mithilfe geeigneter Quellen, tatsächlicher Benutzer oder durch Vertreter der Benutzerinteressen im Entwicklungsprozess erfasst. Das Ergebnis ist eine umfassend dokumentierte Beschreibung des Nutzungskontexts (Burghardt et al., 2011).

In der eigenen Arbeit wurden dafür folgenden Methoden verwendet:

### a. Moderierter Usability Test

Unter einem moderierten Usability Test versteht man das Testen einer Applikation unter einheitlich festgelegten Rahmenbedingungen mit der potenziellen Zielgruppe. Dabei werden die Testpersonen anhand vorgegebener Aufgaben von einem Moderator durch den Test geführt. Während des Tests beobachtet der Moderator die Teilnehmer und zeichnet die Bewegungen und Schritte auf. Zusätzlich wird das sogenannte „laute Denken“ angewendet, bei dem die Testpersonen ihre Gedanken und geplanten Handlungen laut aussprechen, um ihre Vorgehensweise besser nachvollziehbar zu machen. Im Anschluss an den Test erfolgt eine anonyme Umfrage, in der die Teilnehmer eine abschließende Bewertung abgeben können (Uebelnickel et al., 2015).

### b. Externes Benchmarking

Unter einem externen Benchmarking versteht man die Recherche, Aufstellung und Analyse von Wettbewerbern in einem ausgewählten Markt. Ziel der Methode ist es vorhandene Schwachstellen sowie vorhandene Stärken zu identifizieren, um die eigene Marktposition zu verbessern (Burkhardt, 2018).

### c. Customer Journey Map

Unter einer Customer Journey Map versteht man ein Diagramm, das die Schritte, Gedanken, Emotionen und Gefühle von Nutzern, während sie versuchen ein erwünschtes Ziel zu erreichen, in einem Prozess visualisiert. Dadurch lassen sich Probleme und Defizite identifizieren, die noch verbessert oder ergänzt werden müssen (Lewrick, 2020).

## 2. Nutzungsanforderungen spezifizieren

Aus den Benutzeranforderungen sowie den funktionalen, finanziellen und gesetzlichen Vorgaben werden die Anforderungen an die neue Applikation abgeleitet, wobei der zuvor definierte Nutzungskontext berücksichtigt wird (Burghardt et al., 2011).

In der eigenen Arbeit wurden dafür folgende Methoden verwendet:

a. Personas

Die Personas veranschaulichen typische Nutzer einer Zielgruppe und repräsentieren deren Ziele, Wünsche, Bedürfnisse und Ängste, um Entscheidungen bei der Erstellung neuer Produkte und Dienstleistungen treffen zu können. Dabei werden fiktive Charaktere erstellt, die jedoch auf realen Benutzerdaten basieren. Eine Persona fasst die Merkmale mehrerer realer Personen zusammen und erhält zusätzlich einen Namen, ein Geschlecht sowie grundlegende demografische Daten wie Alter, Beruf und Hobbys. Ergänzend werden auch Persönlichkeitseigenschaften und Charaktermerkmale definiert (Lewrick, 2020).

b. User Stories

Eine User Story ist eine einfach formulierte Software-Anforderung, die aus der Perspektive eines typischen Nutzers einer Zielgruppe verfasst wird. Sie beschreibt, welche Nutzergruppe welche Funktionen innerhalb eines Systems benötigt und welchen Nutzen sie daraus zieht (Kauer-Franz & Franz, 2022).

c. How might we...?

Das Ziel der „How might we...?“-Methode („Wie können wir...?“) besteht darin, zuvor definierte Probleme in Fragestellungen zu übersetzen. Die Methode kann dabei helfen über mögliche Lösungen nachzudenken (Uebelnickel et al., 2015).

### 3. Lösungen auf Basis der Nutzungsanforderungen entwickeln

Im nächsten Schritt werden Gestaltungslösungen auf Basis des aktuellen Stands der Technik, der Erfahrungen der Beteiligten und der Ergebnisse der Analyse des Nutzungskontextes entwickelt. Diese Lösungen werden mithilfe von Prototypen konkretisiert (Burghardt et al., 2011).

In der eigenen Arbeit wurden dafür folgende Methoden verwendet:

a. Brainstorming

Unter dem Brainstorming versteht man eine Methode zur Ideenfindung durch das Sammeln von möglichst vielen spontanen Einfällen. Auch ungewöhnliche und unkonventionelle Ideen sind ausdrücklich erwünscht. Es gilt also Quantität vor Qualität. Eine Bewertung der Ideen soll dabei bewusst erst nach der Ideengenerierung durchgeführt werden. Eine gute Grundlage für die Ideenfindung stellt dabei die „How might we...?“-Methode dar (Uebelnickel et al., 2015).

b. Offenes Card Sorting

Unter einem offenen Card Sorting versteht man eine Methode, die dabei hilft, schnell und einfach eine Navigationsstruktur für eine Anwendung zu erstellen, die auf die Anforderungen der Zielgruppe abgestimmt ist. Dabei ordnen die Teilnehmenden die Inhalte der Anwendung in sinnvolle Gruppen, bringen sie in eine logische Reihenfolge

und versehen sie mit passenden Oberkategorien. Zudem haben sie die Möglichkeit, Inhalte umzubenennen, um die Benutzerfreundlichkeit noch weiter zu verbessern (Jacobsen & Meyer, 2022).

c. Future-State Journey Map

Die Future-State Journey Map visualisiert die zukünftigen Schritte der Nutzer in einem Prozess, um ein ideales Nutzungserlebnis darzustellen. Die Methode kann zudem als Leitfaden bei der Entwicklung der neuen Anwendung dienen und hilft, die wichtigsten Funktionen und Inhalte zu priorisieren, um die dargestellte optimale Nutzererfahrung zu realisieren (Kaplan, 2020).

d. Informationsarchitektur

Unter einer Informationsarchitektur versteht man eine strukturierte und organisierte Darstellung aller Seiten und Funktionen innerhalb einer Applikation (Jacobsen & Meyer, 2022).

e. Wireframes

Wireframes sind schnelle, skizzenhafte Darstellungen einer Applikation, die auf funktionale Aspekte beschränkt sind und demnach ohne Farbe oder Design-Details auskommen. Sie werden oft auf Papier oder auch digital erstellt, um erste Ideen und Konzepte zu visualisieren und zu bewerten (Uebelnickel et al., 2015).

f. Styleguide

Ein Styleguide definiert verbindliche Gestaltungsrichtlinien für die Elemente einer Applikation. Er umfasst die Typografie mit Vorgaben zu Schriftart, -größe, Abständen und Verwendungszweck sowie eine Farbpalette mit klaren Regeln zur Anwendung. Zudem legt er die Bildsprache, Icons und deren Einsatz fest. Ergänzend kann ein Gestaltungsraster oder Grid-System enthalten sein (Jacobsen & Meyer, 2022).

g. Low-Fidelity-Prototype

Ein Low-Fidelity-Prototype ist ein erster grober Entwurf einer Applikation ohne viel Funktionen, der genutzt wird, um durch frühe Tests erstes Feedback zu sammeln und die Applikation weiter zu verbessern. Er ist bewusst noch sehr einfach gehalten, ohne Farbe oder detailliertes Design. Der Fokus liegt darauf, grundlegende Konzepte zu testen, bevor Zeit in Feinheiten investiert wird (Jacobsen & Meyer, 2022). Um die Handlungen und Gedankengänge der Testperson besser nachvollziehen, dokumentieren und anschließend auswerten zu können, wird die Methode des sogenannten „lauten Denken“ eingesetzt. Dabei sollen die Gedanken ungefiltert, ungeschönt und offen ausgesprochen werden (Hahn, 2020).

#### **h. High-Fidelity-Prototype**

Ein High-Fidelity-Prototype ist eine weiterentwickelte, detaillierte Version einer Applikation, die auf dem Feedback des Low-Fidelity-Prototyps basiert. Er ist interaktiv, hat ein ausgearbeitetes Design und bildet die wesentlichen Funktionalitäten der Anwendung ab. Ziel ist es, das Nutzungserlebnis und das Grundkonzept vor der eigentlichen Entwicklung umfassend zu testen (Jacobsen & Meyer, 2022). Auch hier soll erneut die Methode des sogenannten „lauten Denken“ angewendet werden, um die Handlungen und Gedankengänge der Testperson besser nachvollziehen, dokumentieren und anschließend auswerten zu können. Dabei sollen die Gedanken ungefiltert, ungeschönt und offen ausgesprochen werden (Hahn, 2020).

### **4. Lösungen gegen Nutzungsanforderungen testen**

In diesem letzten Schritt werden die Prototypen von den Benutzern bewertet. Das erhaltene Nutzerfeedback fließt anschließend in die Überarbeitung ein. Dieser iterative Prozess wird so lange fortgesetzt, bis die benutzerzentrierten Gestaltungsziele vollständig erreicht sind (Burghardt et al., 2011).

In der Arbeit wurden deshalb für jeden der Prototypen moderierte Usability-Tests mit Umfragen durchgeführt.

#### **a. Moderierter Usability-Test**

Unter einem moderierten Usability-Test versteht man das Testen einer Applikation unter einheitlich, festgelegten Rahmenbedingungen mit der potenziellen Zielgruppe. Dabei werden die Testpersonen anhand vorgegebener Aufgaben von einem Moderator durch den Test geführt. Während des Tests beobachtet der Moderator die Teilnehmer und zeichnet die Bewegungen und Schritte auf. Zusätzlich wird das sogenannte „laute Denken“ angewendet, bei dem die Testpersonen ihre Gedanken und geplanten Handlungen laut aussprechen, um ihre Vorgehensweise besser nachvollziehbar zu machen. Im Anschluss an den Test erfolgt eine anonyme Umfrage, in der die Teilnehmer eine abschließende Bewertung abgeben können (Uebenickel et al., 2015).

#### **b. Feedback Capture Grid**

Das Feedback Capture Grid dient dazu, das Feedback der Testpersonen zu erfassen oder zu ordnen. Dabei wird das Feedback in vier Kategorien aufgeteilt: positives Feedback, Änderungsanmerkungen, Fragen und Ideen zur Verbesserung der Applikation (Uebenickel et al., 2015).

## **2.2 Grundlagen KI**

In den folgenden Unterkapiteln wird sowohl auf die Entwicklung und Funktionsweise von KI als auch auf potenzielle Risiken und Gefahren im privaten und beruflichen Umfeld eingegangen sowie auf die Anwendung von KI im Webdesign.

### **2.2.1 Basiswissen KI**

Der Begriff „Künstliche Intelligenz“, kurz KI, fand das erste Mal Erwähnung im Mai 1956 auf der Dartmore Conference am Dartmouth College in New Hampshire (Beranek, 2021). Doch bereits in der Antike befassten sich griechische Philosophen wie Aristoteles mit der Logik, insbesondere mit dem sogenannten logischen Schließen. Dabei stellte sich die Frage, „ob man aus bekannten wahren Aussagen logisch korrekt auf andere wahre Aussagen schließen kann“ (Otte, 2021, S.39). Ein klassisches Beispiel dafür lautet: „Alle Menschen sind sterblich. Alle Griechen sind Menschen. Alle Griechen sind sterblich“ (Otte, 2021, S.39). Im Jahr 1847 überführte der britische Mathematiker George Boole diese Logik in ein formales System, die sogenannte boolesche Algebra, mit der sich logische Aussagen mathematisch durch Gleichungen darstellen lassen. Die technische Umsetzung logischer Operationen begann, als Alan Turing 1936 mit an der Entwicklung der Turing-Maschine arbeitete, die verschlüsselte Nachrichten im zweiten Weltkrieg entschlüsselte. Mit dieser Maschine konnte bewiesen werden, dass sich ein universeller Computer konstruieren lässt. Damit wurde der Grundstein für die technische Entwicklung des Computers gelegt. Konrad Zuse entwickelte daraufhin 1941 die Z3, der erste Universalcomputer. Diese Maschine war jedoch im Gegensatz zu der Mark1, die 1944 in den USA entwickelt wurde, noch nicht rein elektronisch, sondern mit elektromechanischen Relais. Aufbauend auf diesen theoretischen und technischen Entwicklungen wurde schließlich die Grundlage für die Entstehung der Künstlichen Intelligenz geschaffen (Otte, 2021).

Heute versteht man unter Künstlicher Intelligenz eine Software, die nicht deterministische, starre Lösungswege verfolgt, sondern durch Erfahrungen lernt und sich kontinuierlich weiterentwickelt (Beranek, 2021). Um als intelligent zu gelten, muss ein System in der Lage sein, auf Umweltreize angemessen zu reagieren und bei erkannten Fehlern seine Funktionsweise eigenständig anzupassen, um künftig eine angemessene Reaktion zu zeigen. Dies setzt die Fähigkeit zum Denken, zum Wissenserwerb sowie zum Lernen aus Erfahrung voraus. In diesem Zusammenhang wird zwischen schwacher und starker KI unterschieden, wobei der Unterschied vor allem in dem Leistungsumfang und den Fähigkeiten liegt. Die schwache KI kann aus zur Verfügung gestellten Daten selbstständig lernen, auch Machine Learning genannt, und auf dieser Basis auch neue Daten erzeugen, auch Machine Thinking genannt (Otte, 2021). Diese Form der KI ist bereits in vielen Bereichen im Einsatz, wie bei der Sprach- oder Bilderkennungen, bei virtuellen Assistenten wie Siri oder Alexa und bei Gesichtserkennungen, um zum Beispiel das Smartphone zu entsperren. Auch in Internetanwendungen ist die KI zu finden wie bei Chatbots, Spamfiltern, Suchmaschinen oder Empfehlungsalgorithmen von Spotify oder Netflix (Beranek, 2021). Die starke KI hingegen ist „ein gedankliches Konstrukt, dass heute in keiner Weise realistisch ist“ (Otte, 2021, S. 38), da sie über ein menschliches Maß an Bewusstsein und ein eigenes Selbstbewusstsein verfügen müsste. Aktuelle KI-Systeme besitzen jedoch weder einen eigenen Willen noch ein Ich-Bewusstsein oder Emotionen (Otte, 2021).

Wie bereits erläutert, basiert das Lernen der KI auf den Daten, die ihr zur Verfügung gestellt werden. Dies kann jedoch auch Risiken und Probleme mit sich bringen, insbesondere wenn die Datengrundlage Verzerrungen enthält. Ein gutes Beispiel hierfür ist der von Amazon entwickelte Algorithmus zur Bewerberauswahl, der weibliche Bewerbende systematisch benachteiligte. Der Grund dafür war, dass die Daten aus den letzten zehn Jahren gezeigt haben, dass hauptsächlich Männer in den Bewerbungen vertreten waren. Das führte dazu, dass die KI Männer als die besseren Kandidaten eingestuft hat, wodurch diese bevorzugt wurden (Schmidt, 2021).

Auch in der Arbeitswelt wird die Entwicklung der KI mit Sorge beobachtet. Im Rahmen des „ver.di-Innovationsbarometer 2019 - Künstliche Intelligenz“ wurden 990 ver.di-Arbeiternehmervertreter in Aufsichtsräten sowie Betriebs- und Personalräten befragt. Dabei äußerten 66 Prozent der Teilnehmenden die Befürchtung, dass durch den Einsatz von KI Arbeitsplätze verloren gehen könnten, während lediglich 3 Prozent mit einem Zuwachs rechnen. Zudem gehen 52 Prozent davon aus, dass die Arbeitsintensität steigen wird, wohingegen nur 8 Prozent eine Entlastung erwarten. 42 Prozent befürchten eine Zunahme von Störungen im Arbeitsprozess, während 11 Prozent eine Verbesserung vorhersagten. Auch die Entscheidungs- und Handlungsspielräume der Beschäftigten könnten laut 60 Prozent der Befragten schrumpfen; nur 4 Prozent rechnen mit einer Erweiterung. Hinsichtlich der Kompetenz- und Qualifikationsanforderungen erwarten 51 Prozent eine Zunahme, während 23 Prozent eine Abnahme annehmen. Auffällig ist, dass 50 Prozent eine höhere Transparenz des Arbeits- und Leistungsverhaltens vorhersagen, wohingegen nur 13 Prozent eine geringere Transparenz erwarten. Insgesamt zeigt sich, dass die befragten Arbeitnehmervertreter dem Einsatz von KI im Unternehmenskontext überwiegend kritisch gegenüberstehen und erhebliche negative Auswirkungen befürchten (vgl. Zanker et al., 2019).

**Abb. 10 Folgen von Künstlicher Intelligenz für die Arbeitsbedingungen in Unternehmen, in denen damit gearbeitet wird**

„Alles in allem: Wie wirkt sich der Einsatz von KI aktuell Ihrer Meinung nach aus?“

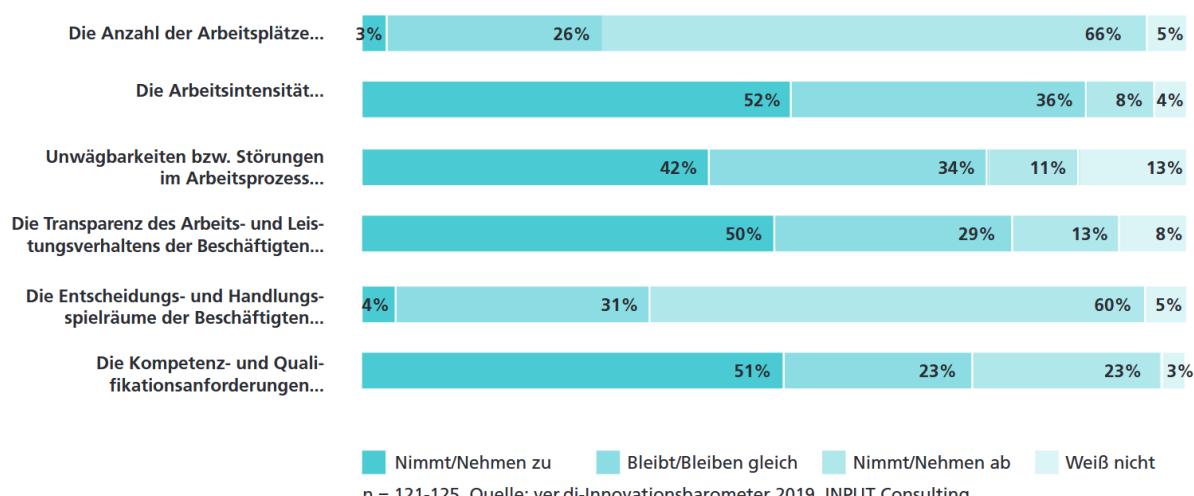


Abbildung 17. Folgen von Künstlicher Intelligenz für die Arbeitsbedingungen in Unternehmen (Zanker et al., 2019, S. 28)

Otte (2021) weist auf weitere Gefahren und Risiken der Künstlichen Intelligenz hin. So warnt er unter anderem vor einer „Verdummung der Menschen“ (Otte, 2021, S. 233) durch den Einsatz von KI in sozialen Medien sowie vor einer potenziellen Entmündigung der Menschen im Kontext KI-gestützter Bildung. Darüber hinaus sieht er die Gefahr einer umfassenden gesellschaftlichen Überwachung als besonders problematisch an und thematisiert auch den Einsatz von KI in militärischen Bereichen als kritischen Aspekt (Otte, 2021, S. 233). Ein großes Risiko besteht zudem in dem unreflektierten Einsatz Künstlicher Intelligenz durch Laien, da der bloße Einsatz intelligenter Systeme keine korrekten Ergebnisse garantiert. Im Gegenteil, die Resultate sind häufig fehleranfällig und benötigen eine kritische Überprüfung (Otte, 2021).

## 2.2.2 KI im Webdesign

Durch das verstärkte Aufkommen von KI wurde das Thema bereits auch im Webdesign Bereich untersucht. So wird in der Veröffentlichung von Mamadaliev Rustambek (2023) mit dem Titel „THE ROLE OF ARTIFICIAL INTELLIGENCE IN WEBSITES“ Einsatzmöglichkeiten, Chancen und Risiken von KI im Bereich Webdesign behandelt. Er beschreibt, wie KI-Algorithmen das Nutzerverhalten beobachten können, um Vorlieben, Verhaltensweisen und vergangene Interaktionen zu analysieren, um anschließend die Inhalte, das Design und die Funktionalität entsprechend anzupassen, um eine verbesserte personalisierte Benutzererfahrung zu ermöglichen. Weiterhin bieten KI-gestützte Chatbots sofortige Unterstützung, können Fragen beantworten und Nutzer durch Prozesse führen, was das Benutzererlebnis insgesamt verbessert. Auch eine automatisierte Erstellung von Inhalten wie Produktbeschreibungen, Nachrichtenartikel und Blogbeiträge auf der Grundlage vordefinierter Parameter und Daten ist möglich. Zudem kann die KI interessante Inhalte aus verschiedenen Quellen sammeln und vorschlagen, die zu den Vorlieben der Website-Besucher passen. Außerdem kann sie helfen, die Sicherheit zu erhöhen, indem sie Bedrohungen frühzeitig erkennt und darauf reagiert. Es werden jedoch auch die Risiken von KI bezüglich Datenschutz- und Sicherheitsbedenken deutlich, da Nutzerdaten gesammelt und verarbeitet werden. Das verantwortungsvolle Nutzen und Einbinden von KI in Applikationen sowie Transparenz und Datenschutz sind aus ethischen Gründen wichtig, um auch das Vertrauen der Nutzer langfristig sicherzustellen. Es wird jedoch deutlich, dass KI zukünftig weiterhin eine zentrale Rolle in der Optimierung von Websites und Apps einnehmen wird.

Biplov Paneru et al (2024) beschreiben zusätzlich noch die Möglichkeit von KI Analysen durchzuführen, die das Benzerverhalten und Vorlieben vorhersagen können, sodass UI/UX-Designer vorausschauend Applikationen entwerfen können, die den Erwartungen der Nutzer entsprechen und somit die Benutzererfahrung verbessern. Zudem kann die KI zu einer Verbesserung der Barrierefreiheit, durch Funktionen wie Bilderkennung, Verarbeitung natürlicher Sprache und Sprachbefehle, beitragen. Auch der Prozess der Evaluierung von UI/UX-Elementen kann mit Hilfe der künstlichen Intelligenz automatisiert werden, indem die KI durch Analysen Probleme in der Nutzerfreundlichkeit identifiziert und Einblicke in das Benutzerverhalten ermöglicht. Dies hilft den Designern die Tests schneller durchzuführen und die Applikation zu verbessern.

Jedoch zeigen auch Arbeiten der Nielsen Norman Group, dass KI-Tools zwar den Design Prozess verbessern und unterstützen können, jedoch zum aktuellen Stand die Designer nicht komplett ersetzen können (Sponheim & Brown, 2024) (Nielsen, 2021).

Eine erste Literaturrecherche zeigt also, dass die Einsatzmöglichkeiten sowie die Chancen und Risiken von KI im Bereich UI/UX bereits gut erforscht sind, dennoch wird ein direkter Vergleich zwischen einer KI erstellten und einer manuell erstellten Applikation interessante Ergebnisse liefern und eine Forschungslücke füllen.

## 2.3 Grundlagen der Datenanalyse

In den folgenden Unterkapiteln wird sowohl auf den Mann-Whitney-U-Test zur Ermittlung der statistischen Signifikanz als auch auf die deskriptive Statistik eingegangen.

### 2.3.1 Mann-Whitney-U-Test

Zur Auswertung der erhobenen Umfragedaten wurde der Mann-Whitney-U-Test angewendet, der es ermöglicht zu überprüfen, ob sich zwei unabhängige Gruppen hinsichtlich einer bestimmten abhängigen Variable statistisch signifikant unterscheiden. Er wird eingesetzt, wenn die Voraussetzungen für den t-Test nicht erfüllt sind, weil die Daten zum Beispiel nicht normalverteilt oder nicht intervallskaliert sind. Statt mit den tatsächlichen Messwerten arbeitet der Test mit Rangplätzen (Tokarski, 2008).

Damit der Mann-Whitney-U-Test angewendet werden kann, müssen folgende Voraussetzungen erfüllt sein:

1. Die Anzahl der getesteten Gruppen beträgt mindestens zwei (Tokarski, 2008).

Im Rahmen der Tests wurden die Testpersonen in zwei Gruppen von jeweils einundzwanzig Personen aufgeteilt, die jeweils eine der beiden Websites testete. Somit stehen für die Auswertung zwei voneinander getrennte Gruppen zur Verfügung, deren Ergebnisse miteinander verglichen werden können.

2. Die Gruppen sind unabhängig voneinander (Tokarski, 2008).

Die beiden Gruppen wurden getrennt gebildet und es gab keine Überschneidungen oder wechselseitige Beeinflussungen. Die Messwerte der einen Gruppe hatten keinen Einfluss auf die der anderen. Da es sich nicht um eine verbundene Messung handelt, ist die Voraussetzung der Unabhängigkeit erfüllt.

3. Die Daten liegen mindestens ordinalskaliert vor (Tokarski, 2008).

Für die Datenerhebung wurde der User Experience Questionnaire (UEQ) verwendet. Die Antwortmöglichkeiten folgen einer geordneten Skala, die, wie für eine Ordinalskala typisch, eine Rangordnung der Bewertungen ermöglicht (Behnke & Behnke, 2006).

4. Die Daten der Gruppen sollten ähnliche Verteilungen haben (Tokarski, 2008).

Die Diagrammansicht von Google Forms zeigt, dass die Daten der beiden Gruppen ähnlich schief verteilt sind. Somit kann von einer vergleichbaren Verteilung ausgegangen werden.

Für die Durchführung des Mann-Whitney-U-Tests müssen die Daten zunächst in eine Rangordnung gebracht werden. Dabei wird jeder Wert einem Rang zugeordnet, unabhängig von der

Gruppenzugehörigkeit. Das heißt, dass alle Werte der beiden Gruppe in eine gemeinsame Rangordnung gebracht werden. Der niedrigste gemessene Wert bekommt demnach den Rangplatz 1, der nächsthöhere Wert den Rangplatz 2 und so weiter. Treten zwei gleiche Werte auf, die sich einen Rangplatz teilen, werden für diese Werte ein mittlerer Rangplatz gebildet. Dieser ergibt sich aus dem Mittelwert der betroffenen Rangplätze. Teilen sich beispielsweise drei Werte die Ränge 2 bis 7, erhält jeder dieser Werte einen Rangplatz von 4,5. Der nächste Wert wiederum erhält den Rangplatz 8. Im Anschluss daran wird die Summe der Rangplätze für jede Gruppe berechnet, die für weitere Berechnungen notwendig ist (Rasch et al., 2014).

Formel zur Berechnung der Rangsumme einer Gruppe (Rasch et al., 2014):

$$T_i = \sum_{m=i}^{n_i} R_{mi}$$

Die beteiligten Gruppen werden mit dem Index  $i$  bezeichnet. Die Anzahl der Werte innerhalb einer Gruppe  $i$  wird mit  $n_i$  angegeben. Jeder einzelne Wert innerhalb einer Gruppe wird dabei einem Rangplatz  $R_{mi}$  zugewiesen, wobei  $m$  für den  $m$ -ten Wert innerhalb der Gruppe  $i$  steht (Rasch et al., 2014).

Nach der Berechnung der Rangsummen für beide Gruppen kann bereits ein Unterschied festgestellt werden. Jedoch ist nicht erkennbar, ob der Unterschied statistisch signifikant ist. Dies kann mithilfe der Summe der Rangplatzüberschreitung  $U$  festgestellt werden, die zeigt, wie oft ein Rangplatz aus der einen Gruppe größer ist als ein Rangplatz aus der anderen Gruppe. Sie gibt also die Summe der Rangplatzüberschreitungen der einen Gruppe gegenüber der anderen Gruppe an (Rasch et al., 2014).

Formel zur Berechnung der Summe der Rangplatzüberschreitungen der ersten Gruppe (Rasch et al., 2014):

$$U_1 = n_1 \times n_2 + \frac{n_1 \times (n_1 + 1)}{2} - T_1$$

Analog dazu die Formel zur Berechnung der Summe der Rangplatzüberschreitungen der zweiten Gruppe (Rasch et al., 2014):

$$U_2 = n_1 \times n_2 + \frac{n_2 \times (n_2 + 1)}{2} - T_2$$

Die Anzahl der Werte aus der Gruppe 1 wird mit  $n_1$  und aus der Gruppe 2 mit  $n_2$  angegeben. Die Summe der Rangplätze, die den Versuchspersonen aus Gruppe 1 zugeordnet wurden, wird als  $T_1$  oder als  $T_2$  für die zweite Gruppe angegeben (Rasch et al., 2014).

Die Signifikanz des U-Tests wird unter der Annahme der Nullhypothese geprüft. Diese besagt, dass es keinen bedeutsamen Unterschied zwischen den beiden Gruppen in Bezug auf die Anzahl der Rangplatzüberschreitungen und -unterschreitungen gibt. Die Wahrscheinlichkeit des tatsächlich beobachteten Ergebnisses kann über die Verteilung der U-Werte unter der Annahme der Nullhypothese bestimmt werden (Rasch et al., 2014).

Die Beziehung von  $U_1$  und  $U_2$  kann mit Hilfe folgender Formel dargestellt werden (Rasch et al., 2014):

$$U_1 = n_1 \times n_2 - U_2$$

Die Werte  $U_1$  und  $U_2$  stehen demnach in einer komplementären Beziehung zueinander. Das heißt, wenn es viele Rangüberschreitungen zwischen den beiden Gruppen gibt, gibt es automatisch weniger Rangunterschreitungen und umgekehrt. Das bedeutet: Je größer der Unterschied zwischen den Gruppen, desto mehr Rangüberschreitungen und weniger Rangunterscheidungen gibt es oder umgekehrt und desto stärker unterscheiden sich  $U_1$  und  $U_2$ . Gibt es hingegen keinen Unterschied zwischen den Gruppen, sind Rangüber- und Rangunterschreitungen gleich verteilt. Die Nullhypothese des U-Tests geht deshalb davon aus, dass  $U_1$  gleich  $U_2$  ist. Wenn die Anzahl der Rangplatzüberschreitungen deutlich größer oder kleiner ist als die der Rangplatzunterschreitungen, spricht das für eine ungleiche Verteilung der Rangplätze zwischen den beiden untersuchten Gruppen. Das kann ein Hinweis darauf sein, dass sich die Gruppen in Bezug auf das untersuchte Merkmal unterscheiden. Geht man jedoch von der Nullhypothese aus, lässt sich daraus ein erwartender Mittelwert für  $U$  berechnen (Rasch et al., 2014).

Formel für den erwartete Mittelwert für  $U$   $\mu_U$  (Rasch et al., 2014):

$$\mu_U = \frac{n_1 \times n_2}{2}$$

Wenn der  $U$ -Wert deutlich vom erwarteten Mittelwert  $\mu_U$  abweicht, egal ob in positiver oder negativer Richtung, ist das ein Hinweis gegen die Nullhypothese. Die Verteilung der  $U$ -Werte um  $\mu_U$  ist symmetrisch. Liegen die Rangüberschreitungen auf der einen Seite, befinden sich die Rangunterschreitungen auf der anderen. Dadurch sind  $U_1$  gleich  $U_2$  auch immer gleich weit vom Mittelwert  $\mu_U$  entfernt (Rasch et al., 2014).

Die zentrale Frage ist, wie stark der beobachtete  $U$ -Wert vom unter der Nullhypothese erwarteten Mittelwert  $\mu_U$  abweichen muss, damit der Unterschied als statistisch signifikant gilt. Um dies zu prüfen, wird eine sogenannte Kennwerteverteilung durchgeführt. Dabei sind alle möglichen  $U_1$ - und  $U_2$  - Werte symmetrisch um den erwarteten Mittelwert  $\mu_U$  angeordnet. Die Streuung der  $U$ -Werte, also wie stark die Werte um den Mittelwert  $\mu_U$  schwanken, wird durch folgende Formel unter Berücksichtigung der verbundenen Ränge berechnet (Rasch et al., 2014):

$$\sigma_{U_{corr}} = \sqrt{\frac{n_1 \times n_2}{N \times (N - 1)}} \times \sqrt{\frac{N^3 - N}{12} - \sum_{i=1}^k \frac{t_i^3 - t_i}{12}}$$

$N$  wird aus der Summe der Werte  $n_1$  und  $n_2$  berechnet. Die Anzahl der Personen, die sich den Rangplatz  $i$  teilen wird mi  $t_i$ , die Anzahl der verbundenen Ränge wird mit  $k$  angegeben (Rasch et al., 2014).

Wenn  $n_1$  oder  $n_2$  größer als 20 ist und ihre Größen nicht zu stark voneinander abweichen, dann verhält sich die Verteilung der  $U$ -Werte ungefähr wie eine Normalverteilung. Dadurch ist es möglich die Standardnormalverteilung zur Prüfung zu verwenden. Der beobachtete  $U$ -Wert, der Mittelwert  $\mu_U$  und die Streuung  $\sigma_{U_{corr}}$  werden in die folgende Formel zur Berechnung eines  $z$ -Werts eingesetzt (Rasch et al., 2014):

$$z_U = \frac{U \times \mu_U}{\sigma_U}$$

Anschließend wird mithilfe des ermittelten p-Wert angegeben, wie wahrscheinlich das Ergebnis unter der Annahme ist, dass die Nullhypothese stimmt. Er hilft dabei zu entscheiden, ob ein statistisches Ergebnis signifikant ist. Ist der p-Wert kleiner als das zuvor festgelegte Signifikanzniveau von  $\alpha = 0,05$ , wird die Nullhypothese abgelehnt. Ist der p-Wert größer als  $\alpha = 0,05$  ist der Unterschied zwischen den beiden Gruppen nicht statistisch signifikant. Der p-Wert wird mit Hilfe des zuvor ermittelnden z-Wert in einer Tabelle der Standardnormalverteilung abgelesen. Für einen zweiseitigen Test „ist der Flächenanteil, welcher durch die Prüfgröße abgeschnitten wird, zu verdoppeln“ (Bortz & Schuster, 2010, S.107). Der p-Wert wird mit folgender Formel bestimmt (Bortz & Schuster, 2010):

$$P(z) = P(z) \times 2$$

Für einen negativen z-Wert ergibt sich folgende Formel (Bortz & Schuster, 2010):

$$P(-z) = (1 - P(z)) \times 2$$

Wie in Abbildung 18 zu erkennen ist, kann beispielsweise für einen z-Wert von -0,11 der Wert 0,5438 abgelesen werden. Die Berechnung des p-Wertes unter Berücksichtigung des negativen z-Wertes und des zweiseitigen Tests ergibt:

$$P(-0,11) = (1 - 0,5438) \times 2 = 0,9124$$

Da der Wert  $p = 0,9124$  das Signifikanzniveau von  $\alpha = 0,05$  deutlich übersteigt, kann die Nullhypothese nicht abgelehnt werden, wodurch der Unterschied nicht statistisch signifikant ist.

z-Wert*	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,00	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,10	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,20	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,30	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,40	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,50	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,60	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,70	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,80	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,90	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389

Abbildung 18. Ausschnitt der Tabelle zur Verteilungsfunktion der Standardnormalverteilung (Bortz & Schuster, 2010, S. 587)

Dabei ist anzumerken, dass die meisten Tabellen der Standardnormalverteilung z-Werte nur bis  $z = 3,89$  aufführen. Muss jedoch ein z-Wert größer als 3,89 berücksichtigt werden, kann mit einem Wert von  $P(z) = 0,9999$  gerechnet werden. Wie in Abbildung 19 erkennbar ist, verändert sich der p-Wert ab  $z = 3,63$  nur noch in der fünften Nachkommastelle. Eine fünfte Nachkommastelle und damit ein genauerer Wert von  $P(z)$  wird das Ergebnis der statistischen Signifikanz nicht beeinflussen.

3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999

Abbildung 19. Ausschnitt der Tabelle zur Verteilungsfunktion der Standardnormalverteilung für einen hohen z-Wert (Uni Hamburg, n.d.)

Mithilfe der Größe eines Effekts  $r$  kann anschließend die Teststärke bestimmt werden. Also ein Maß für die Stärke eines Zusammenhangs, eines Unterschieds zwischen zwei Gruppen oder Effekts in der Statistik (Bortz & Schuster, 2010). J. Cohen (1988) hat als Orientierungspunkte für die Effektstärke folgende Werte festgelegt:

Tabelle 1. Werte der Effektstärke  $r$  nach Cohen (1988)

Effektstärke	Interpretation
$r = 0,2$	kleiner Effekt
$r = 0,5$	mittelgroßer Effekt
$r = 0,8$	großer Effekt

Als möglichen Grund für eine kleine Effektgröße sieht Cohen (1988) die fehlende genaue Kontrolle über die Bedingungen oder Messungen während der Tests und der Folge, dass unkontrollierbare Einflüsse wie Geräusche als Störfaktoren das Ergebnis verfälschen. Ein mittelgroßer Effekt sieht Cohen (1988) bereits als groß genug an, um bereits vor dem Test und der statistischen Auswertung erkannt werden zu können. Er führt dabei den Unterschied des IQs von akademischen und leitenden Berufsgruppen auf, der im normalen Alltag bereits erkennbar ist. Ein großer Effekt besagt, dass die zwei Gruppen so stark voneinander getrennt sind, dass ihre Werte sich fast nicht überlappen.

Bei dem nichtparametrischen Mann-Whitney-U-Test kann die Effektstärke mithilfe des z-Werts und der Stichprobengröße  $n$  berechnet werden (Fritz et al., 2012):

$$r = \frac{|z|}{\sqrt{n}}$$

### 2.3.2 Deskriptive Statistik

Für spätere Berechnungen muss zunächst der Mittelwert, auch als arithmetisches Mittel oder Durchschnitt bezeichnet, bestimmt werden. In der Statistik stellt er ein Maß zur Beschreibung der zentralen Tendenz einer Verteilung dar. Er wird berechnet, indem die Summe aller Werte durch die Anzahl der Werte geteilt wird.

Formel zum Berechnen des Mittelwertes (Bortz & Schuster, 2010):

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Der Mittelwert ist gegenüber jedem einzelnen Wert der Stichprobe sensitiv, das heißt jede Veränderung eines Wertes beeinflusst den Mittelwert. Dies macht ihn zu einem präzisen Maß für das Zentrum einer Verteilung, aber auch anfällig für Ausreißer oder Extremwerte. Einzelne ungewöhnlich hohe oder niedrige Wert können den Mittelwert erheblich verschieben und somit eine Verzerrung der zentralen Tendenz erzeugen. Solche Ausreißer können durch Datenfehler oder Berechnungsfehler entstehen. Dadurch ist es ratsam, zusätzlich den Median zu betrachten, da dieser robuster gegenüber Ausreißern ist und damit ein stabileres Maß für die zentrale Lage der Daten bietet (Bortz & Schuster, 2010).

Der Median stellt einen Wert dar, der genau in der Mitte einer nach Größe sortierten Datenreihe liegt. Das bedeutet, dass 50 % der Werte kleiner und 50 % größer als dieser Wert sind. Für die Ermittlung des Medians werden die Werte zuerst nach ihrer Größe sortiert. Die sogenannte Ordnungsstatistik wird wie folgt beschrieben (Bortz & Schuster, 2010):

$$x_{(1)}, x_{(2)}, \dots, x_{(n)}.$$

Dabei stellt  $x_{(1)}$  den kleinsten Wert und  $x_{(n)}$  den größten Wert dar.

Der Median kann anschließend mit Hilfe folgender Formel bestimmt werden (Bortz & Schuster, 2010):

$$Md = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & \text{für } n \text{ ungerade} \\ (x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}) \div 2 & \text{für } n \text{ gerade} \end{cases}$$

Das Ergebnis der Formel gibt anschließend die Position des Medians  $Md = x_{(i)}$  in der sortierten Datenreihe an. Der Median ist robust gegen Ausreißer, da selbst wenn sehr große oder sehr kleine Werte in der Stichprobe stark verändert werden, der Median meist unverändert bleibt. Wenn befürchtet wird, dass ungewöhnlich oder extreme Werte in den Daten vorkommen, ist es sinnvoll, den Median statt des Mittelwertes als Maß für die zentrale Tendenz zu nutzen (Bortz & Schuster, 2010). Die Verwendung des Medians ist zudem bei einer Ordinalskala, wie sie der UEQ aufweist, empfehlenswert, da diese die Daten in einer Reihenfolge liefert und dies beim Ermitteln des Medians benötigt wird. Die Berechnung des Mittelwertes ist dennoch zusätzlich sinnvoll, weil er eine einfache und aussagekräftige Kennzahl darstellt, um den durchschnittlichen Wert einer Verteilung zu beschreiben. Er ermöglicht es, verschiedene Gruppen einer Umfrage schnell miteinander zu vergleichen, ohne die gesamte Verteilung betrachten zu müssen (Schäfer, 2016).

Um die Streuung der Werte zu betrachten kann der Interquartilsabstand verwendet werden, der robust gegenüber Ausreißern ist und dennoch eine präzise Streuung der Daten wiedergibt. Dafür werden die

Daten wie bei der Ermittlung des Medians in eine nach Größe sortierte Reihenfolge gebracht, die anschließend in vier gleich große Teile aufgeteilt werden, die Quartile genannt werden. Anschließend wird die Differenz aus dem oberen und dem unteren Quartil berechnet. Um den unteren Quartil, auch 25% Quartil genannt, zu bestimmen wird der mittlere Wert des unteren Datensatzes bestimmt. Besteht die Hälfte aus einer geraden Anzahl von Werten wird der Mittelwert aus den mittleren Werten bestimmt. Sollte der gesamte Datensatz eine ungerade Anzahl haben erhält sowohl die untere Hälfte als auch die obere Hälfte den Wert, der in der Mitte steht. Um das obere Quartil, auch 75% Quartil genannt, zu berechnen wird der mittlere Wert der oberen Hälfte des Datensatzes ermittelt. Anschließend wird der untere Quartil vom oberen Quartil abgezogen, wodurch man den Interquartilsabstand erhält (Schäfer, 2016).

Formel zu der Bestimmung des Interquartilsabstand beziehungsweise des Abstands zwischen dem 25%-Quantil und dem 75%-Quantil (Behnke & Behnke, 2006):

$$\text{Interquartilsabstand} = x_{0,75} - x_{0,25}$$

Für die Darstellung der Ergebnisse wird das sogenannte Boxplot-Diagramm verwendet. Es kann verwendet werden, um Daten, die zuvor in eine Rangliste gebracht wurden zu veranschaulichen. Dargestellt wird dabei der Median, der Mittelwert, der Interquartilsabstand, die Spannweite also die Gesamtigkeit der Daten und die Ausreißer (Homrighausen, 2020).

## Begriffe der Skala

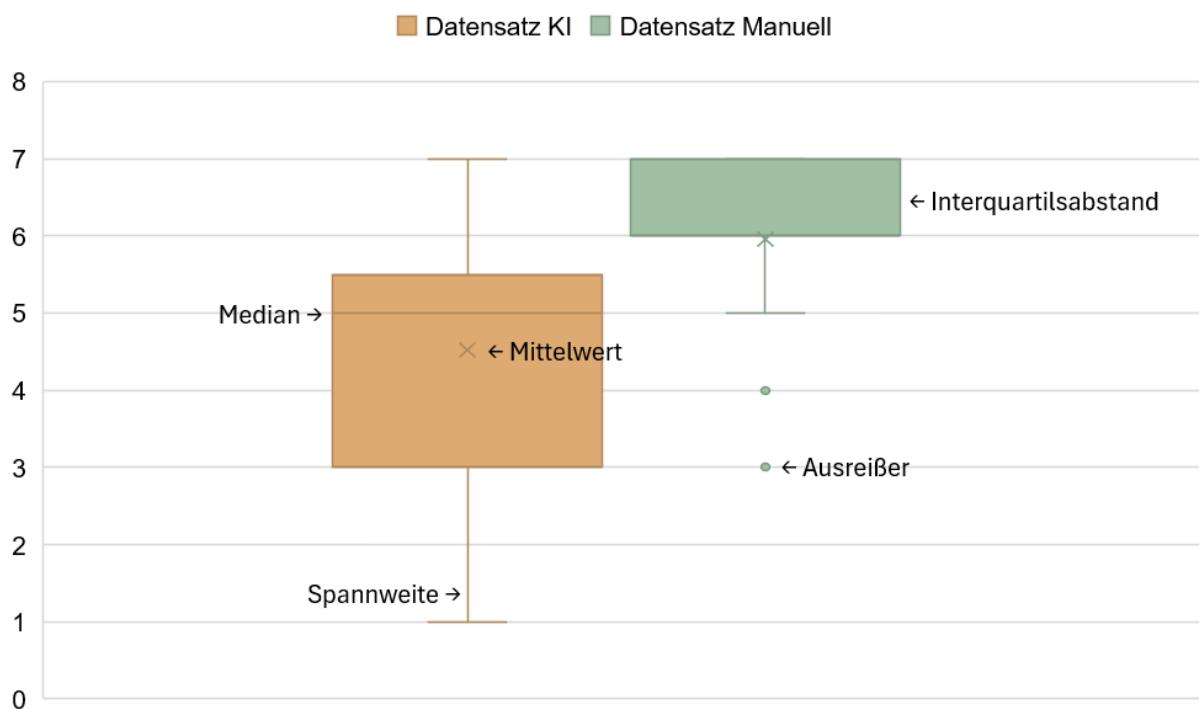


Abbildung 20. Veranschaulichung der Inhalte des Boxplot-Diagramms

## **3. Methodik**

In den folgenden Unterkapiteln wird sowohl die Erstellung der manuell erstellten als auch der KI-erstellten Applikation beschrieben, mit denen später die Tests zur Beantwortung der Forschungsfragen durchgeführt werden.

Zunächst wird die Erstellung der manuell erstellten Applikation dargestellt, die von mir entwickelt wurde. Anschließend wird die Erstellung der KI-generierten Applikation vorgestellt, die mit dem Programm WIX und den Anweisungen einer Person ohne Vorkenntnisse im Webdesign entwickelt wurde.

### **3.1 Erstellung der manuell gestalteten Applikation**

Im Folgenden werden die durchgeführten Methoden des Human-Centered Design-Prozesses und die anschließende Erstellung der Applikation beschrieben.

#### **3.1.1 Verstehen und Beschreiben des Nutzungskontextes**

Mithilfe der in den folgenden Unterkapiteln beschriebenen Methoden wurde der Nutzungskontext auf Grundlage tatsächlicher Nutzer analysiert und definiert (Burghardt et al., 2011).

##### **3.1.1.1 Externes Benchmarking**

Im ersten Schritt des Benchmarkings wurde ein passendes Template erstellt. Anschließend folgte eine Recherche zu möglichen Mitbewerbern, um einen Überblick zu erhalten. Dabei wurden vier relevante Vergleichspartner identifiziert: Tee Kontor Kiel, The English Tearoom, Tee Gschwendner und Königsmann Teeladen. Diese wurden anhand ihrer Ähnlichkeit und Vergleichbarkeit untereinander ausgewählt. Die Mitbewerber wurden nach einer vertieften spezialisierten Recherche in den Überkategorien der Datengewinnung, der Problemraumerkennung, der Problemlösung der Mitbewerber, der User Rezensionen und der Verarbeitung der Ergebnisse untersucht.

Zunächst wurden potenzielle Mitbewerber recherchiert. Dabei wurden sowohl regionale Teeläden aus Erlangen als auch große, bekannte Online-Tee-Shops untersucht. Anschließend wurden diese anhand bestimmter Kriterien bewertet, um als Vergleichspartner in Frage zu kommen. Dies beinhaltete die Untersuchung welche Ähnlichkeiten und Vergleichbarkeiten gegeben sein müssen, um als Vergleichspartner gelten zu können. Die gefunden Ergebnisse waren dabei das Vorhanden sein eines Teeverkaufs von spezielleren Tees, die über das klassische Angebot hinausgehen und nicht nur aus Beuteltees bestehen, ein Angebot an Zubehör und ein Verkauf sowohl vor Ort als auch online. Nicht alle recherchierten Mitbewerber erfüllten diese Anforderungen, sodass die vier oben genannten Vergleichspartner übrigblieben.

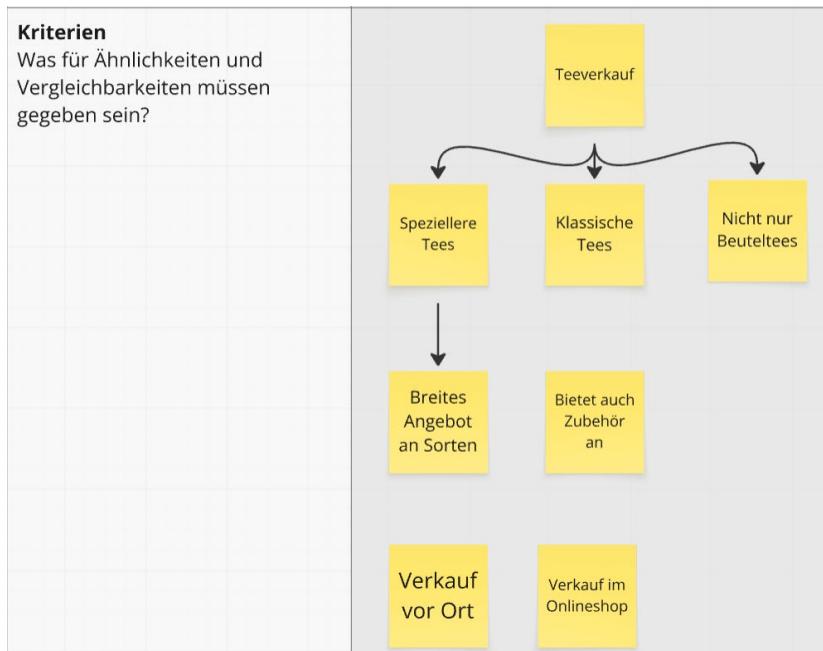


Abbildung 21. Benchmarking – Kriterien der Vergleichspartner

Weiterhin wurde untersucht, ob ein derartiges Problem bereits betrachtet wurde. Hierbei zeigte sich, dass es kaum Informationen dazu gibt, ob die Ladenbesitzer ihre Websites regelmäßig analysieren, testen und optimieren. Es kann jedoch angenommen werden, dass dies individuell unterschiedlich und der Erfolg nicht immer gegeben ist.

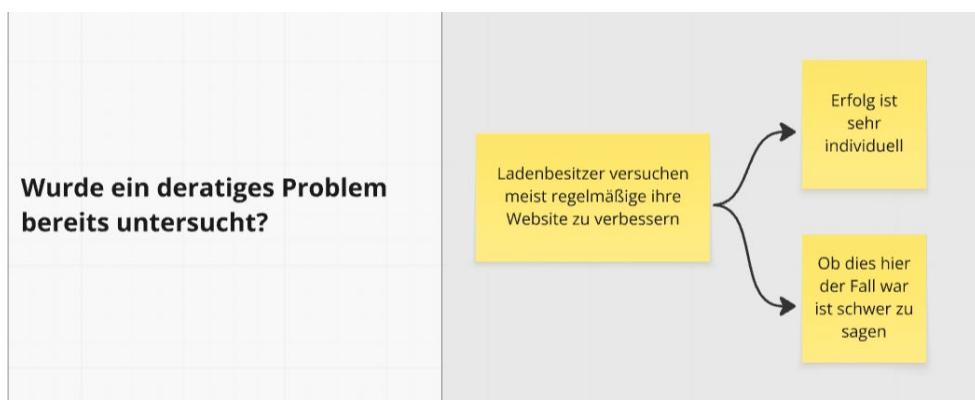


Abbildung 22. Benchmarking – Wurde das Problem bereits untersucht?

Als nächstes wurde in der Kategorie der Datengewinnung die Websites von Tee Kontor Kiel, The English Tearoom, Tee Gschwendner und Königsmann Teeladen genauer untersucht und die wichtigsten Inhalte, Schritte und Funktionen notiert.

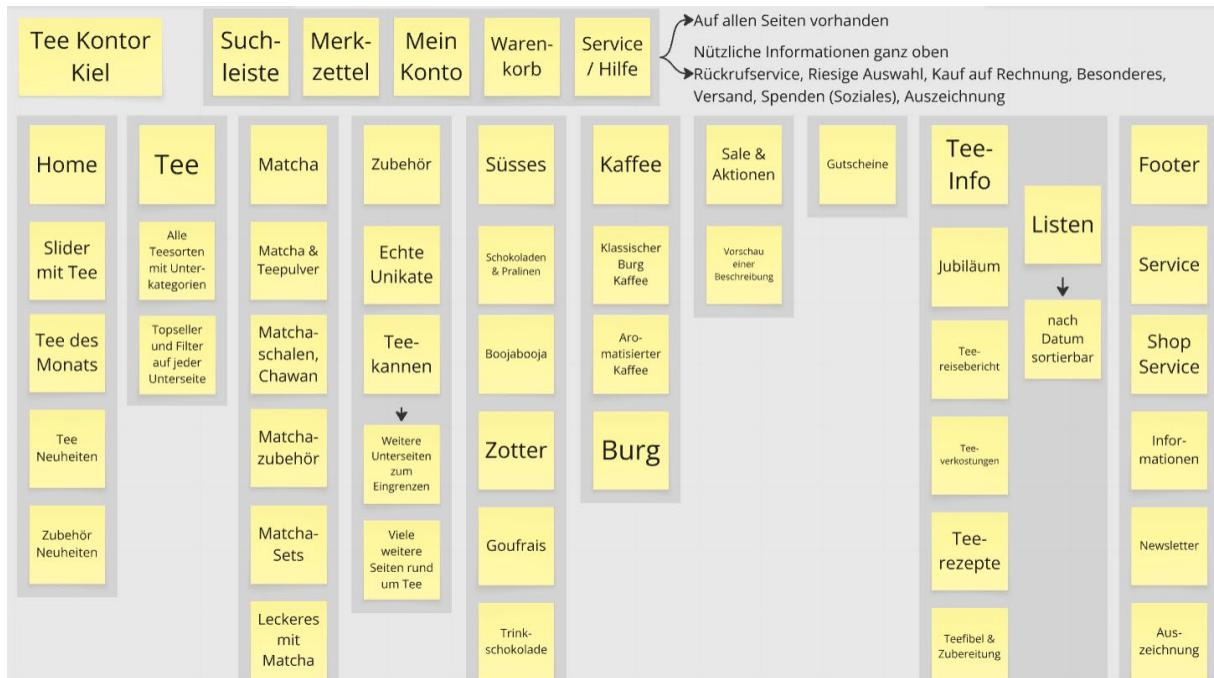


Abbildung 23. Benchmarking – Datengewinnung für die Website von Tee Kontor Kiel



Abbildung 24. Benchmarking – Datengewinnung für die Website von Königsmann Teeladen



Abbildung 25. Benchmarking – Datengewinnung für die Website von Tee Gschwendner

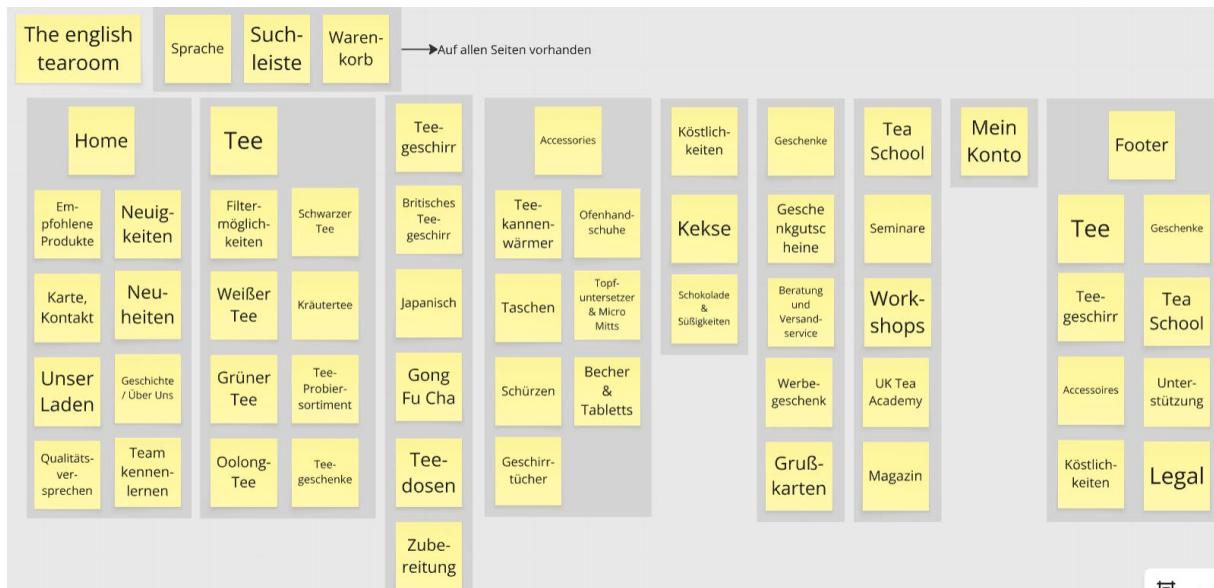


Abbildung 26. Benchmarking – Datengewinnung für die Website von The English Tearoom

Zudem wurde die Zielgruppe der Websites ermittelt. Dabei ergaben sich zum einen die erfahrenen Teetrinker, die bereits ein umfangreiches Wissen über Tee haben und gezielt bestimmte Sorten suchen. Zum anderen die unerfahrenen Teetrinker, die sich über verschiedene Tees informieren und neue

Sorten ausprobieren möchten. Bezuglich der Altersgruppe ist es zunächst möglich, dass Tee alle Altersgruppen anspricht. Jedoch gilt ab sieben Jahren eine beschränkte Geschäftsfähigkeit<sup>11</sup> (*Bürgerliches Gesetzbuch (BGB) § 106 Beschränkte Geschäftsfähigkeit Minderjähriger*, n.d.) mit der Zustimmung des gesetzlichen Vertreters meist in Form der Eltern<sup>12</sup> (*Bürgerliches Gesetzbuch (BGB) § 107 Einwilligung des gesetzlichen Vertreters*, n.d.) und erst ab erreichen der Volljährigkeit mit achtzehn Jahren gilt eine unbeschränkte Geschäftsfähigkeit<sup>13</sup> (*Geschäfts- Und Deliktsfähigkeit*, n.d.). Zudem ist anzunehmen, dass Personen ab dem siebzigsten Lebensjahr eher nur im Geschäft vor Ort einkaufen und kaum mehr online. Dadurch ergibt sich eine Altersgruppe von achtzehn bis siebzig Jahren.

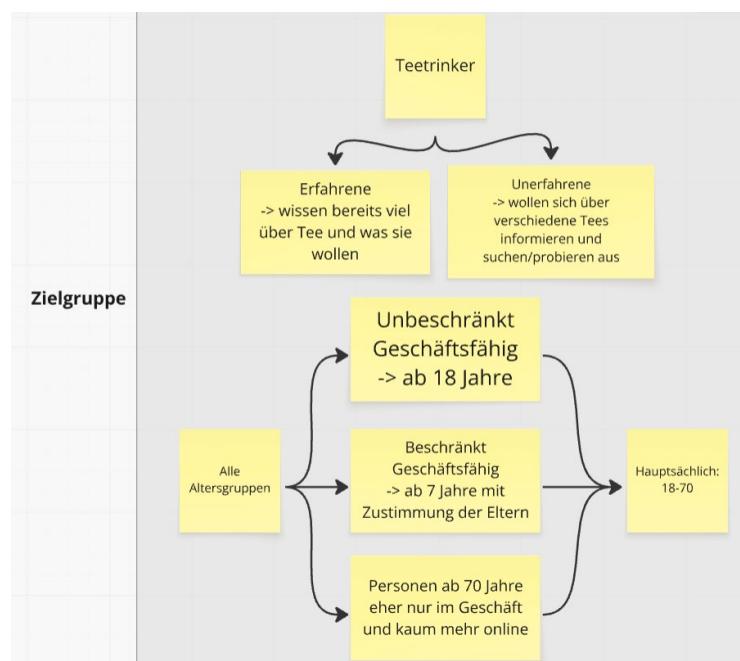


Abbildung 27. Benchmarking – Zielgruppe

Anschließend wurde zu der Kategorie der Problemraumerkennung übergegangen. Dabei wurden zunächst die Stärken jeder Mitbewerber festgehalten. Für die Website von Tee Kontor Kiel war dies die Auffindbarkeit von Topseller und die Filter auf jeder Unterseite, der Jump Up Button, ein einheitliches Design, kleine Erklärtexte mit Informationen auf den Unterseiten, hilfreiche Video Verlinkungen rund um das Thema Tee, die Möglichkeit nach Familien oder Anbauern zu sortieren, gute Filtermöglichkeiten, eine kleine Anzeige wie viel es bereits kostet ist immer sichtbar, Neuheiten und beliebte Produkte sind schnell auf Home auffindbar und nützliche Informationen wie der Rückrufservice, die riesige Auswahl,

<sup>11</sup> [https://www.gesetze-im-internet.de/bgb/\\_106.html](https://www.gesetze-im-internet.de/bgb/_106.html)

<sup>12</sup> [https://www.gesetze-im-internet.de/bgb/\\_107.html](https://www.gesetze-im-internet.de/bgb/_107.html)

<sup>13</sup> <https://www.uni-potsdam.de/de/rechtskunde-online/rechtsgebiete/zivilrecht/geschaefts-und-deliktsfaehigkeit>

Kauf auf Rechnung, Besonderes, Versand, Spenden (Soziales) und Auszeichnungen sind immer ganz oben zu sehen. Auch Inhalte wie die Suchleiste, der Merkzettel, das Kundenkonto und der Warenkorb sind immer präsent und schnell auffindbar im Menü enthalten.

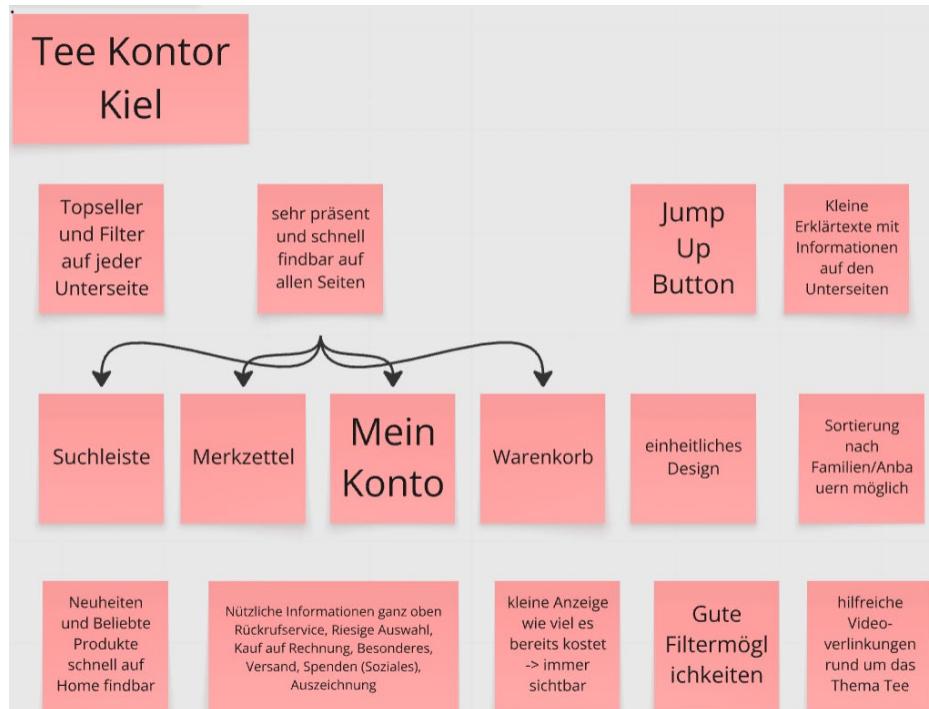


Abbildung 28. Benchmarking – Ermittelte Stärken von Tee Kontor Kiel

Auch die Website von Tee Gschwendner konnte einige Stärken vorweisen, wie kleine Erklärtexte mit Informationen auf den Unterseiten, eine ordentliche und übersichtliche Menüleiste, ein einheitliches Design, Filter auf jeder Unterseite und ein Jump Up Button. Aber auch das cleane Design von Icons, die gut auffindbar sind, wie Standort, Sprache, Suche, Merkzettel, Konto und Warenkorb auf jeder Seite sind als Stärke hervorzuheben. Zudem gibt es nützliche Informationen über den Versand, das Widerufsrecht, ein sicheres Einkaufen und zu PAYPACK.



Abbildung 29. Benchmarking – Ermittelte Stärken von Tee Gschwendner

Die Stärken der Website The English Tearoom waren zum einen das einheitliche Design, keine überfüllten, sondern aufgeräumte Seiten, eine Suchfunktion und gut auffindbare Inhalte durch eine gute Navigation und Sortierung. Zum anderen gibt es kleine Erklärtexte mit Informationen auf den Unterseiten und auch ein Filtern nach genaueren Kategorien auf jeder Unterseite ist möglich.

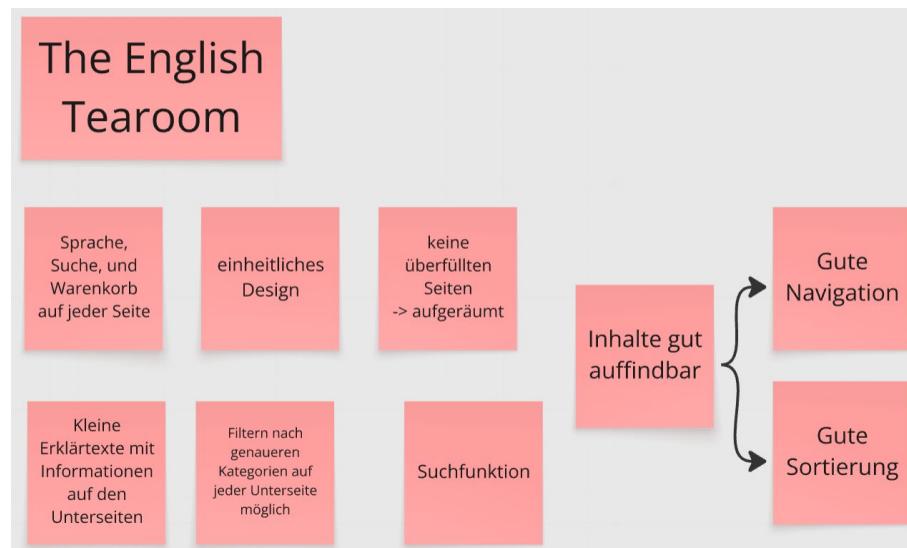


Abbildung 30. Benchmarking – Ermittelte Stärken von The English Tearoom

Für die Website von Königsmanns Teeladen konnte nur die Stärke ermittelt werden, dass wichtige Informationen und Empfehlungen direkt unter Home schnell auffindbar sind.



Abbildung 31. Benchmarking – Ermittelte Stärken von Königsmann Teeladen

Darüber hinaus wurden auch die Schwächen der Mitbewerber untersucht. Die Schwächen der Website von Tee Kontor Kiel sind zum einen, dass die Seiten und der Header sehr voll sind, dies macht beides sehr unübersichtlich. Auch in der Menüleiste, wie zum Beispiel bei Tee oder Zubehör, befindet sich sehr viel Inhalt. Das Menü wird dadurch sehr unübersichtlich und lange überladene Listen sind mit zeitaufwändigen Suchen des gewünschten Produktes verbunden. Zum anderen stehen die Informationen über den Laden erst im Footer und sind damit sehr schwer auffindbar.

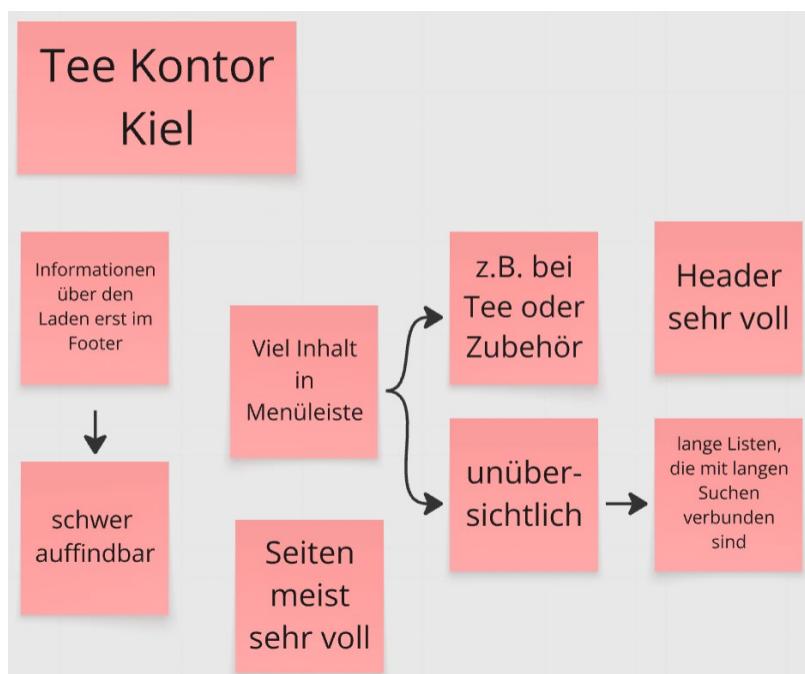


Abbildung 32. Benchmarking – Ermittelte Schwächen von Tee Kontor Kiel

Auch die Website von Tee Gschwendner hatte ein paar Schwächen. Auffallend war hier zum einen das sehr kühle und unpersönliche Design und die fehlenden Informationen über die genaue Herkunft des Tees. Auch die nicht anklickbaren Pfeile bei den Filtermöglichkeiten verwirren.



Abbildung 33. Benchmarking – Ermittelte Schwächen von Tee Gschwendner

Die Website von The English Tearoom hat ebenfalls einige Schwächen. Zum einen sind Inhalte wie Sprache, Suche, und Warenkorb etwas zu klein, die Schrift auf den Bildern in Home ist schlecht lesbar und eher störend, in Home muss lange gescrollt werden und es gibt keinen Jump Up Button. Zum anderen werden nicht alle Inhalte nach dem Umstellen der Sprache in Deutsch übersetzt, die Farben sind sehr grell, die Website wirkt unruhig, da es sehr viele Muster und verschiedene Effekte gibt und unter dem Header wurde nicht einheitlich auf allen Seiten ein großes Bild platziert, das jedoch auch einen großen Teil der Seite einnimmt. Zuletzt sind die Informationen zu dem Laden nicht präsent genug, sondern sehr auf Home verstreut und die Aktion „Add to cart“ ist in der Übersicht aller Produkte verwirrend, da man zu der Unterseite des ausgewählten Produktes weitergeleitet wird, um das Produkt dort in den Warenkorb zu legen.

## The english tearoom

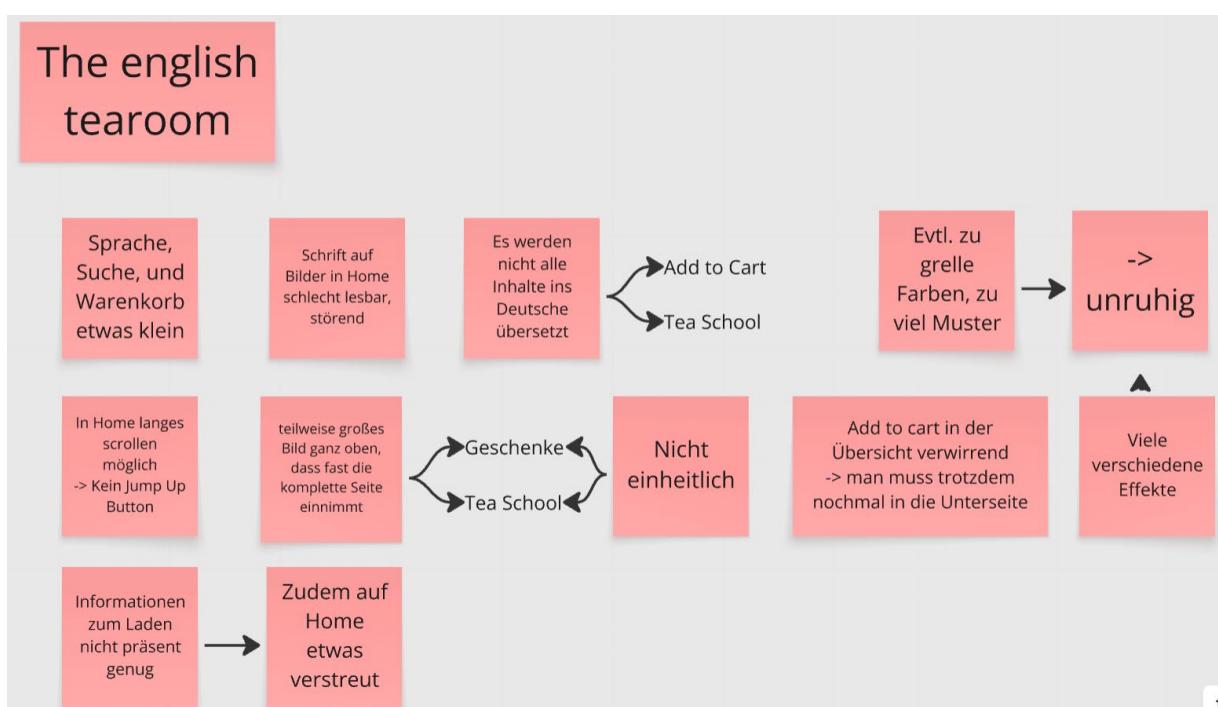


Abbildung 34. Benchmarking – Ermittelte Schwächen von The English Tearoom

Auch die Website von Königsmanns Teeladen wies Schwächen auf. Zum einen gibt es zwei Menüs und keinen klassischen Footer, da dessen Inhalte in einem Burger Menü versteckt sind. Zum anderen ist das Menü am linken Rand nicht fest, sondern verschwindet beim Scrollen und in Home ist ein langes Scrollen ohne einen Jump Up Button möglich. Weiterhin gibt es viel Fließtext auf der Startseite, die gelbe Schrift auf dem dunkelorange Hintergrund ist nicht gut lesbar und es ist keine Suchfunktion vorhanden. Zudem gibt es lange Listen, die unübersichtlich sind und langes Suchen und Durchklicken nach sich ziehen. Zuletzt gibt es kaum Unterseiten, sondern meistens öffnen sich Pop-ups.

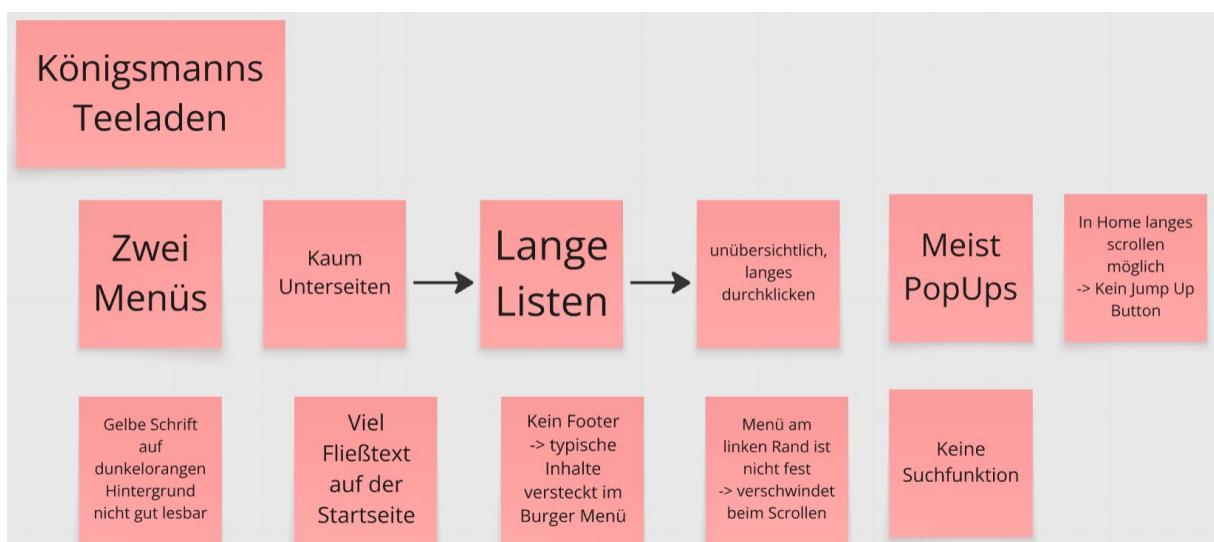


Abbildung 35. Benchmarking – Ermittelte Schwächen von Königsmann Teeladen

Darüber hinaus wurden markante Unterschiede der Websites herausgearbeitet. Diese bestanden vor allem in dem Design, den Inhalten des Headers und der Darstellung der Informationen über den Laden,

da die Priorisierung des Online-Shops oder des Ladens unterschiedlich waren. Auch die Darstellung von Inhalten in Unterseiten oder Pop-ups war unterschiedlich.

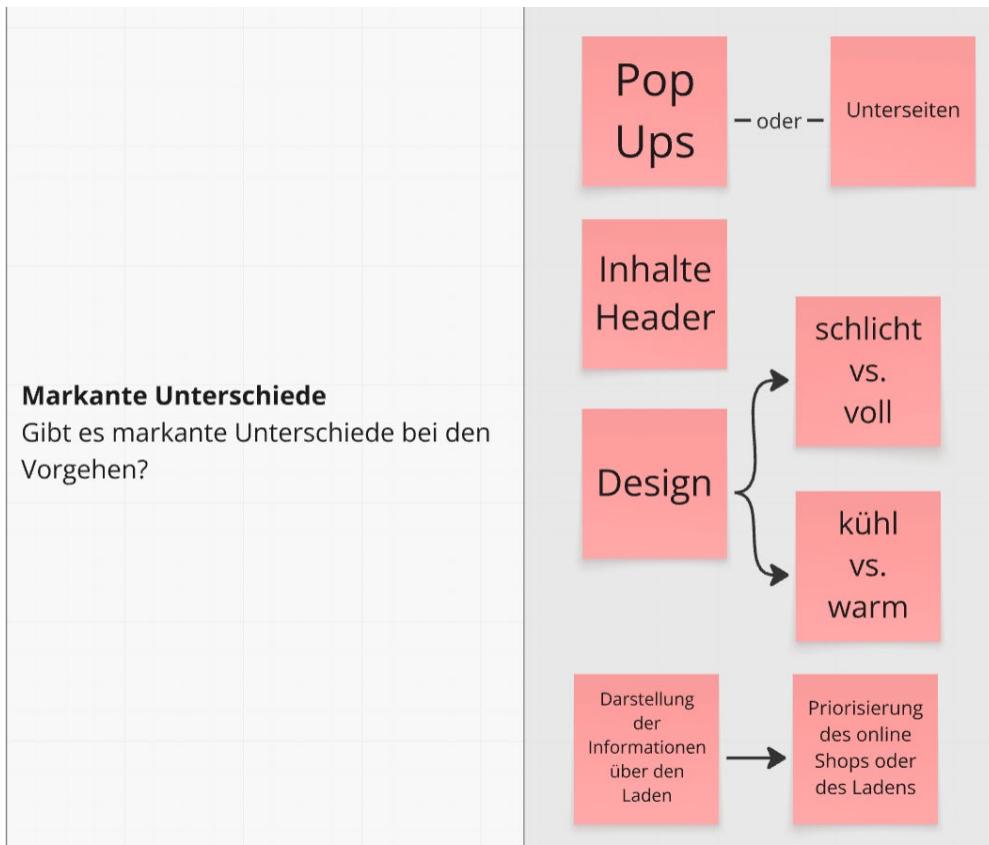


Abbildung 36. Benchmarking – Ermittelte markante Unterschiede

In der nächsten Kategorie der Problemlösung wurde analysiert, ob die Mittbewerber wirklich das vermeintliche Problem lösen. Für die Website von Königsmanns Teeladen ist diese Frage klar mit nein zu beantworten. Für die anderen Websites ist diese Frage nicht klar beantwortbar, da die Websites zwar zum Teil bereits sehr gut entwickelt sind jedoch noch weiter verbessert werden könnten.

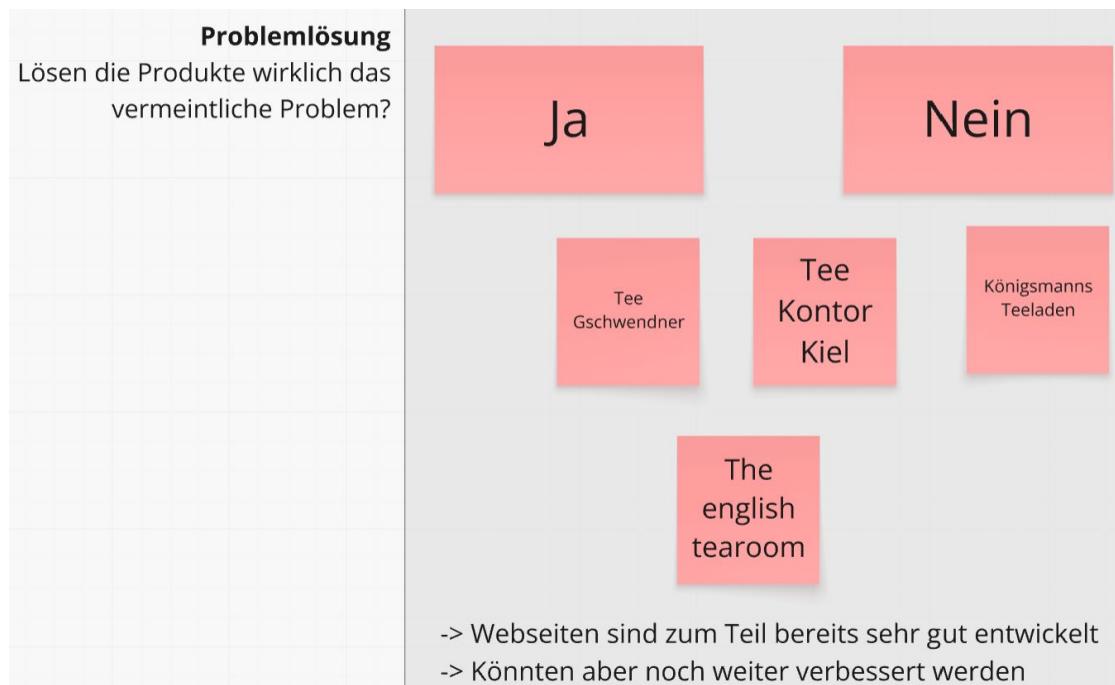


Abbildung 37. Benchmarking – Lösen die Produkte das Problem?

In der Kategorie der User Rezensionen wurde untersucht, welche Bewertungen die Nutzer zu den Websites abgegeben. Verwendet wurde dabei, wenn vorhanden das Kundenfeedback von der eigenen Website sowie von Google. Für den Mitbewerber Tee Gschwendner gab es auf der eigenen Seite Feedback der Kunden, die alle fünf von fünf Sterne vergaben und hauptsächlich die gute Qualität, die Lieferung, die Verpackung und den Service loben. Jedoch ist zu erwähnen, dass es sich zwar um verifizierte Bewertungen handelt es aber nicht deutlich ist woher die Bewertungen stammen und ob negative Bewertungen eventuell einfach nicht angezeigt werden. Auf Google erhält Tee Gschwendner 4,6 Sterne bei 199 Rezensionen. Die positiven Bewertungen loben vor allem die Auswahl an Tees und Zubehör, die Beratung und die Qualität des Tees. Negative Bewertungen bemängeln die Ablehnung von Bargeld und die Kartenzahlung als einzige Bezahlmöglichkeit. Ältere Bewertungen kritisieren jedoch auch das Personal als unprofessionell und unhöflich. Bewertungen, die sich explizit auf die Website beziehen, konnten nicht gefunden werden.

Der Teeladen Königsmanns Teeladen erhielt 4,8 Sterne bei 33 Rezessionen. Positiv wird der Service, die kompetente Beratung, der leckere Tee und die schnelle Lieferung erwähnt. Negativ wird jedoch die geringe Auswahl und der wenig einladende Verkaufsraum angegeben. Einige Kunden geben an auch regelmäßig den Online-Shop zu nutzen.

The English Tearoom erhält 4,9 Sterne bei 176 Rezensionen. Positiv wird hier der Teeladen, das Produktangebot und der Service beschrieben. Eine Kundin gab an, sich vorher auf der Website über den Laden informiert zu haben und eine weitere Kundin berichtete, dass die Website sehr praktisch sei.

Tee Kontor Kiel erhielt ebenfalls 4,9 Sterne bei 124 Rezensionen. Hier wird die Website und der Online-Shop öfter erwähnt und als gut aufgebaut und informativ bewertet. Ansonsten wird der schnelle Versand, der Service und das Angebot gelobt.

In der letzten Kategorie der Methode wurden die Ergebnisse verarbeitet. Zunächst wurde zusammengefasst, welche Punkte bei der eigenen Website umgesetzt werden müssen. Die wichtigsten Punkte waren hierbei eine gute Filtermöglichkeit, eine ordentliche und übersichtliche Menüleiste und Unterseiten, ein einheitliches und cleanes Design, eine gute Beschreibung des Tees und dessen Zubereitung sowie dessen Herkunft.

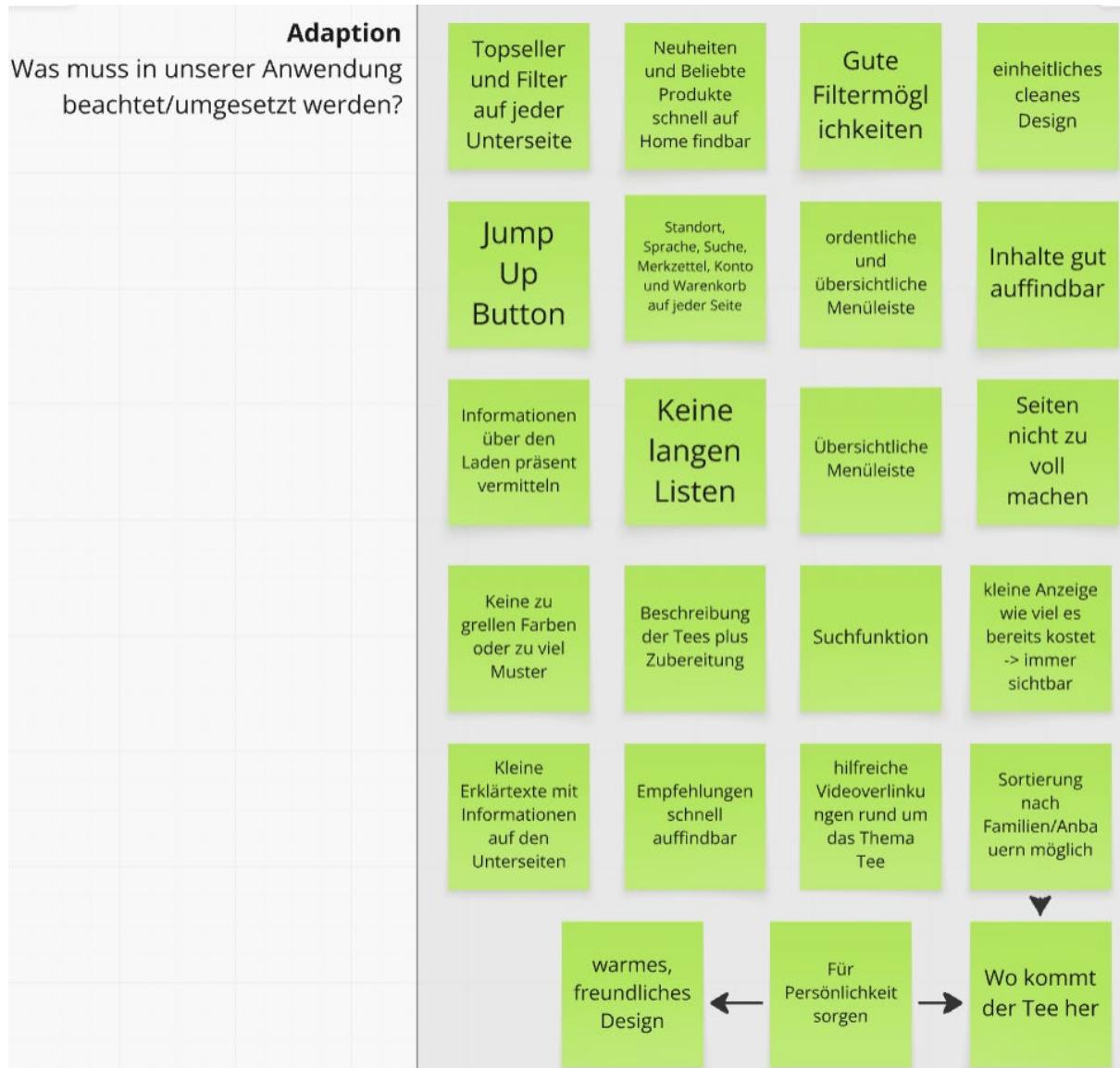


Abbildung 38. Benchmarking – Was muss in der eigenen Anwendung umgesetzt werden?

Als letzter Schritt wurden die Erkenntnisse der Methode zusammengetragen. Insgesamt lieferte die Wettbewerbsanalyse wertvolle Einblicke in die Stärken und Schwächen der Mitbewerber, woraufhin sich viele Chancen und Verbesserungsmöglichkeiten für die eigene Website entwickeln können. Zudem konnte eine gute Übersicht über die notwendigen Inhalte und deren sinnvolle Anordnung erzielt werden.

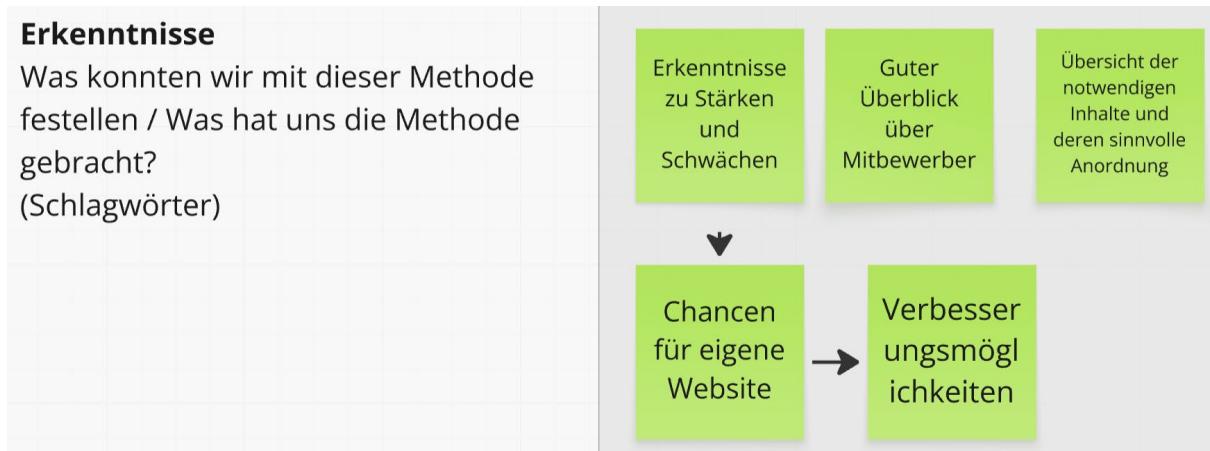


Abbildung 39. Benchmarking – Zusammenfassung der Erkenntnisse der Methode

### 3.1.1.2 Tests bestehender Anwendungen

Zunächst wurden zwei Mitbewerber aus der vorherigen Methode Benchmarking ausgewählt. Die Wahl fiel auf die Websites von Tee Kontor Kiel und The English Tearoom, da sie inhaltlich der eigenen Webseite am ähnlichsten sind und eine gute Grundlage für die Tests darstellten.

Anschließend wurde ein detaillierter Ablaufplan für die Tests erstellt. Dieser umfasst eine kurze Einführung in das Projekt, eine Erklärung des Testablaufs, der Test der ersten Website, die erste Umfrage, der Test der zweiten Website, die zweite Umfrage sowie einen Abschluss mit Danksagung. Gleichzeitig wurden passende Aufgaben für die Tests entwickelt, die am besten dazu geeignet waren, die Benutzerfreundlichkeit und -erfahrung abzutesten. Zudem wurden passende Fragen formuliert und die Umfrage mit dem Tool Google Formulare erstellt.

Die Tests wurden mit insgesamt fünf Testpersonen durchgeführt, darunter zwei Experten in dem Bereich Webdesign und drei Laien. Die Usability Tests wurden moderiert durchgeführt, das heißt, die Testpersonen erhielten die Aufgaben schrittweise und wurden so durch den Test geführt. Während der Tests wurde jeweils die Zeit gemessen und Notizen angefertigt. Dabei wurde besonders auf die Zufriedenheit der Teilnehmenden, dem erfolgreichen Abschluss der Aufgaben sowie der Interaktion mit der Navigation und dem Interface geachtet. Die anschließenden Umfragen wurden anonym beantwortet. Um Verzerrungen durch eine mögliche Lernkurve zu vermeiden, wurde die Reihenfolge der Tests für die Teilnehmenden variiert.

Die gewonnenen Erkenntnisse aus den Usability Tests und den Umfragen wurden in zwei verschiedenen Feedback Capture Grids für jeweils eine der beiden Websites festgehalten.

### 3.1.1.3 Feedback Capture Grid

Zunächst wurde in Miro eine Vorlage für die Methode erstellt. Für die Website von Tee Kontor Kiel und The English Tearoom wurde jeweils ein Feedback Capture Grid angelegt und anhand der Testergebnisse ausgefüllt. Nach Abschluss der Tests der bestehenden Anwendungen wurden die Mitschriften durchgegangen und die Beobachtungen sowie Anmerkungen der Testpersonen erfasst. Auch die Umfrageergebnisse wurden ausgewertet und in die Methode integriert.

Bei der Website von Tee Kontor Kiel wurde als positiv die ansprechende und simple Schriftart, die angenehmen Farben, die Suchfunktion, die Weiterleitung zu Home nach Klick auf das Logo und die große Auswahl bewertet. Auch das Auffinden der Tees hat durch eine gute Produktsortierung und -aufteilung gut funktioniert. Das Finden von Informationen, insbesondere über Tee, fiel den Testpersonen leicht und die YouTube Videos die alles tiefergehend erklären sowie die Definition der Unterkategorien und Unterseiten wurden als nützlich bewertet. Auch bezüglich des Warenkorbs wurden positive Punkte gefunden. Zum einen der Warenkorb Pop-up, das klassische Aussehen des Warenkorbs der bekannt ist und zum anderen die Funktion die Versandkosten zu berechnen. Zuletzt wurde die Möglichkeit Probe- packungen zu bestellen als positiv bewertet.



Abbildung 40. Feedback Capture Grid für Tee Kontor Kiel – What worked

Jedoch ergaben sich auch einige Punkte, die nicht gut funktioniert haben und in der eigenen Website verbesserten werden müssen. Zum einen ist das Menü und der Footer sehr voll und unübersichtlich. Zum anderen ist das Footer Menü in dessen Unterseiten nochmal am linken Rand zu sehen, was von den Testpersonen als unnötig und verwirrend empfunden wurde. Weiterhin wurde bei dem Menü bemängelt, dass dies schnell wieder verschwindet, wenn man leicht mit der Maus aus dem Menü raus geht. Dies störte die Testpersonen vor allem bei den großen Menüs. Zudem empfand auch ein Teil der Tester die Definitionen der Kategorien und der Unterseiten als unnötig und den Warenkorb Pop-up als zu voll. Die zwei „Zur Kasse“ Buttons im Warenkorb verwirrten die Testpersonen sehr. Auch das Design wurde als veraltet und nicht einheitlich empfunden. Zudem gibt es zu viele verschiedene Effekte und auch das Vergrößern der Produktbilder beim drüber hovern wurde als störend empfunden. Weiterhin ist die Website zum Teil nicht responsive angepasst und die Fenster, die in den Unterseiten scrollbar sind,

stören. Auch bei dem Suchen nach Tee wurden ein paar negative Punkte angemerkt. Zum einen war die Kennzeichnung der Filter auf den Teeunterseiten nicht gut genug gekennzeichnet, das Aussehen der Bilder war nicht immer ansprechend, es gab zu viel Fließtext und die Möglichkeit die Tees zu vergleichen wurde nicht gesehen. Zuletzt wurde die Auffindbarkeit der Standortinformationen und Informationen über den Laden als schwer empfunden.

What could be improved						
Zu viele verschiedene Effekte	Menü voll und unübersichtlich	Footer voll und unübersichtlich	Inhalte über der Suche unnötig und störend	Veraltetes Design	Zurückgehen erleichtern -> Von einem Produkt zur Übersichtsseite zurück	
Filter auf Teeunterseiten besser kennzeichnen	Nach Firma oder Herkunft?	zwei zur Kasse Buttons beim Warenkorb	Footer Menü in den Unterseiten nochmal am linken Rand -> unnötig, verwirrend	Definition der Kategorien/Unterseiten bei z.B. Sale und Aktion unnötig	Menü verschwindet leicht wieder, wenn man mit der Maus etwas raus kommt -> stört vor allem bei großen Menüs	Kein einheitliches Design
Vergleichmöglichkeit wurde nicht gesehen	Aussehen der Bilder	Weniger Fließtext	Warenkorb Popup etwas voll	Fenster in den Unterseiten die scrollbar sind	Größe und Formatierung im Menü nicht gleich	Home sehr voll -> wenig Abstand zu Inhalten, unseriöse und störende Inhalte
			Vergrößern der Produktbilder beim drüber hoovern	Zum Teil nicht responsive angepasst	Warnmeldung über Versandkosten erschreckt	Auffindbarkeit der Standortinformationen bzw. Über Uns

Abbildung 41. Feedback Capture Grid für Tee Kontor Kiel – What could be improved

Jedoch ergaben sich auch einige Ideen und Wünsche der Testpersonen. Zum einen wurde vorgeschlagen die Anzeige der YouTube Videos in der Website durch eine einfache Verlinkung nach YouTube zu ersetzen, um den Inhalt zu reduzieren und die Website übersichtlicher zu machen. Zum anderen sollte es eine Verlinkung zu einer eigenen Über Uns Seite ganz oben in Home oder in der Menüleiste geben. Um die Bilder ansprechender zu gestalten könnte man zudem hauptsächlich Bilder der Verpackung verwenden. Im Warenkorb könnte man weitere Produkte anbieten und verlinken, die man noch kaufen könnte, um kostenlosen Versand zu erreichen. Zuletzt sollte das X zum Entfernen eines Produktes durch einen Mülleimer ersetzt werden, da dies verständlicher ist.



Abbildung 42. Feedback Capture Grid für Tee Kontor Kiel – Ideas

Auch für die Website von The English Tearoom ergaben sich sowohl positive als auch einige negative Punkte und Ideen. Als positiv wurde zum einen das moderne und aktuelle Design genannt, das zudem auch eine gute Unterscheidung der verschiedenen Elemente ermöglicht, da diese sich gut farbig voneinander abheben. Auch die ansprechenden Icons im Footer, die großen und schönen Buttons und das Design der Standortkarte wurden als positiv bewertet. Weiterhin wurden das übersichtliche Menü, die einfache Navigation und der Pfad in den Unterseiten, an dem man sich orientieren kann und auch zurück gehen kann, gelobt. Zum anderen fiel das Auffinden des Tees den Testpersonen leicht, da die Sortierung gut ist und es praktische Filtermöglichkeiten gibt. Die Auswahl der Produkte ist zudem immer weiter einschränkbar, von allgemein bis spezifisch. Zusätzlich gibt es interessante Informationen zu dem Tee, wie zum Beispiel das Öko-Siegel, die Bilder des Tees und die Produktbilder allgemein, die Informationen über das Aussehen und den Inhalt geben und die Verlinkungen zu noch genaueren Informationen. Zuletzt wurde der Warenkorb als bekannt und klassisch bewertet und das Berechnen der Versandkosten gefiel den Testpersonen.



## What worked

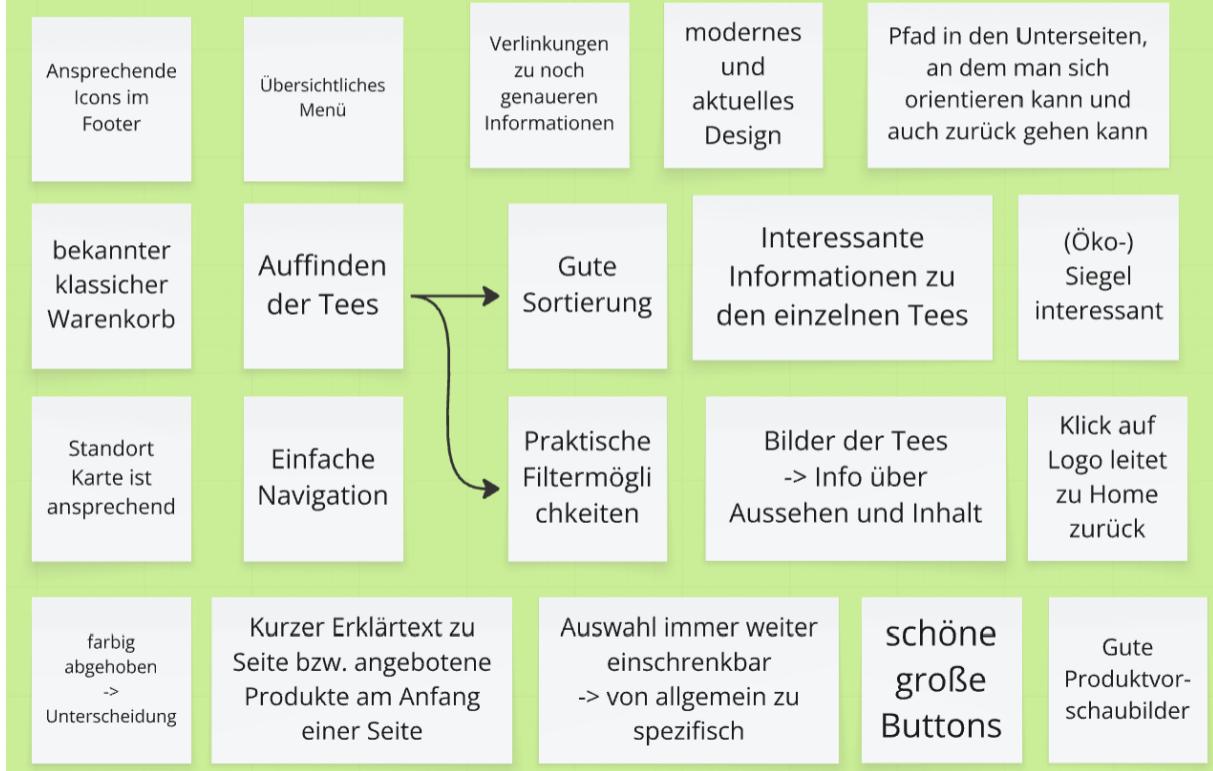


Abbildung 43. Feedback Capture Grid für The English Tearoom – What worked

Jedoch ergaben sich auch hier einige Punkte, die nicht gut funktioniert haben und in der eigenen Website verbessert werden müssen. Zum einen gaben auch Testpersonen an, dass sie die Farben als zu grell und penetrant empfanden und es zu viele verschiedene Farben gibt, deren Bedeutung sie nicht verstanden. Zudem erinnern die Farben eher an einen Süßigkeiten Laden und die Website sowie die Verpackung wirkte „trendy“. Dies hatte zur Folge, dass die Website unpersönlich wirkt. Auch die vielen Effekte bei den einzelnen Elementen empfanden die Tester als störend. Weiterhin gefiel das Aussehen der Bilder nicht und vor allem die Bilder auf Home mit der schlecht lesbaren Schrift waren unpassend. Zusätzlich wurde die Formatierung der Bilder und der Schrift im Warenkorb kritisiert, da diese nicht bündig waren. Zum anderen gab es auch einige negative Anmerkungen bezüglich der Navigation. Die Ladeninformationen waren schwer zu finden, da diese sehr verstreut und nicht klar auffindbar waren. Auch die Teedosen waren aufgrund der schlechten Benennung der Kategorie schlecht zu finden. Kritisiert wurde auch der fehlende Menüpunkt Home. Weiterhin wäre die Sortierung von Matcha unter grünem Tee besser gewesen. Zudem wurden auch die langen Seiten mit viel Inhalt kritisiert, da dies nicht gut für den Gesamtüberblick ist. Weiterhin gab es zu wenig Informationen bezüglich des Verkaufsangebotes im Ladengeschäft, der Versandkosten im Warenkorb und zu wenig Informationen und Erklärungen für Tee Anfänger. Andererseits waren die Informationen zu den einzelnen Kategorien meist zu lang. Die Testpersonen bemängelten auch die fehlenden Kundenbewertungen, den Hinweis auf weitere Seiten unten als zu klein, den Button zur Umstellung der Sprache als zu leicht zu übersehen und den Hinweis oben, dass das Produkt zum Warenkorb hinzugefügt wurde zu als nicht ausreichend. Weiterhin

wurde die Verlinkung zu Google Maps kritisiert, da man ohne erneuter Abfrage sofort weitergleitet wird. Zuletzt wurde der „Add to cart“ Button kritisiert, da dieser nicht auf Deutsch übersetzt wird und von der Übersicht nochmal auf die Unterseite weiterleitet, anstatt das Produkt direkt zum Warenkorb hinzuzufügen.

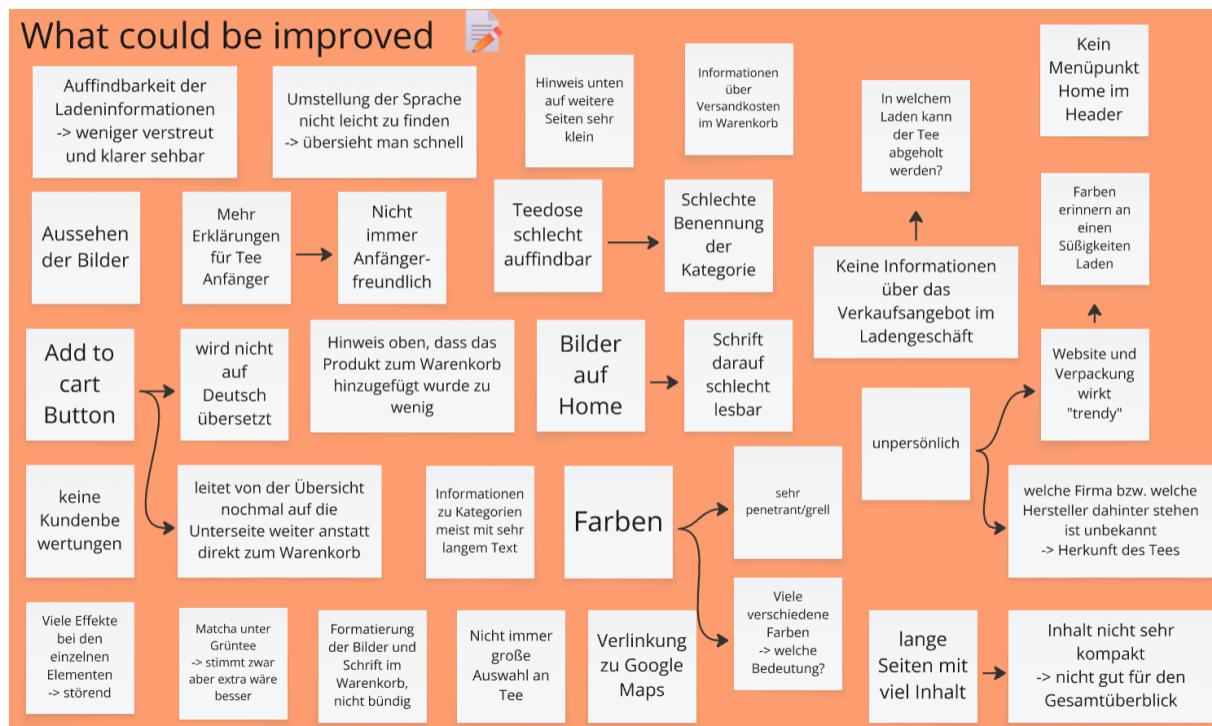


Abbildung 44. Feedback Capture Grid für The English Tearoom – What could be improved

Es gab jedoch auch einige Ideen der Testpersonen, wie man die Website verbessern könnte. Zum einen könnte man eine Vergleichsmöglichkeit der Tees einrichten, Kundenbewertungen anzeigen und hauptsächlich Bilder der Tees mit Verpackung verwenden, um diese ansprechender zu machen. Zum anderen könnte man eine Abfrage vor der Weiterleitung bei Verlinkungen hinzufügen und mehr Informationen über das Verkaufsangebot im Ladengeschäft geben.

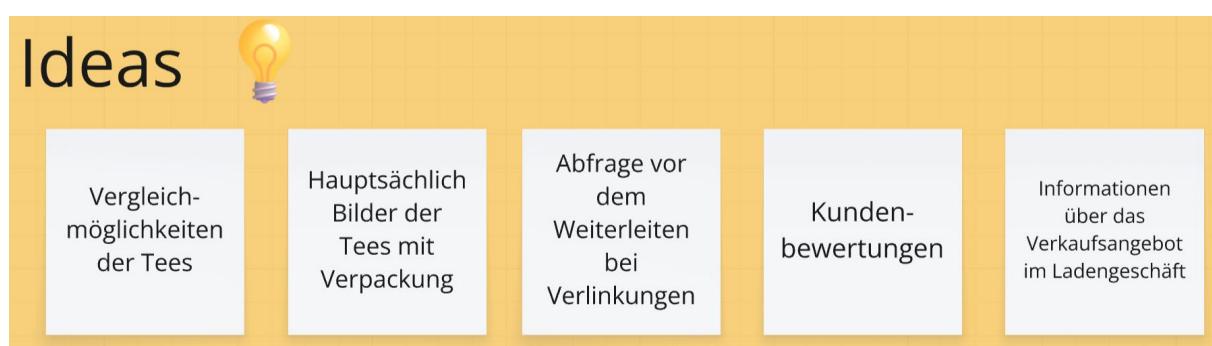


Abbildung 45. Feedback Capture Grid für The English Tearoom – Ideas

Mithilfe dieser Methode konnte das Feedback der Testpersonen übersichtlich festgehalten werden, wodurch eine klare Grundlage und Übersicht für die weiteren Methoden geschaffen wurde.

### 3.1.1.4 Customer Journey Map

Im ersten Schritt wurde eine passende Vorlage für die Methode in Miro erstellt. Anschließend wurde festgelegt, für welche Nutzergruppe die Customer Journey Maps angelegt werden sollen. Für die Websites von Tee Kontor Kiel und The English Tearoom wurden jeweils zwei Maps angelegt, eine für einen erfahrenen und eine für einen unerfahrenen Teetrinker. Die Daten beruhten dabei auf den Ergebnissen der Usability Tests.

Zunächst wurde die Customer Journey Map für die Website von Tee Kontor Kiel und den erfahrenen Teetrinker erstellt, der auf der Suche nach neuen Teesorten und Verkäufern ist. Dazu wurde ein Szenario entwickelt, in dem der Kunde schnell eine gute Übersicht über das gesamte Sortiment finden möchte ohne lange nach spezifischen Tees wie japanischem Grüntee zu suchen. Im nächsten Schritt wurde der Kaufprozess in die einzelnen Schritte unterteilt und in der Map festgehalten. Die wesentlichen Schritte waren dabei die Suche in Google, der Überblick über das Sortiment, das Finden der Ladeninformationen, das Aussuchen des Tees, der Kauf des Tees und der Abschluss des Kaufes. Bei dieser Map wurde deutlich, dass es einem erfahrenen Teetrinker zwar möglich ist sich schnell einen Überblick über die verschiedenen Teesorten zu machen und einen geeigneten Tee zu finden, jedoch gibt es vor allem Schwierigkeiten bezüglich des Auffindens von Informationen zu dem Laden. Zudem wurde das überladene Design der Website als hinderlich empfunden.

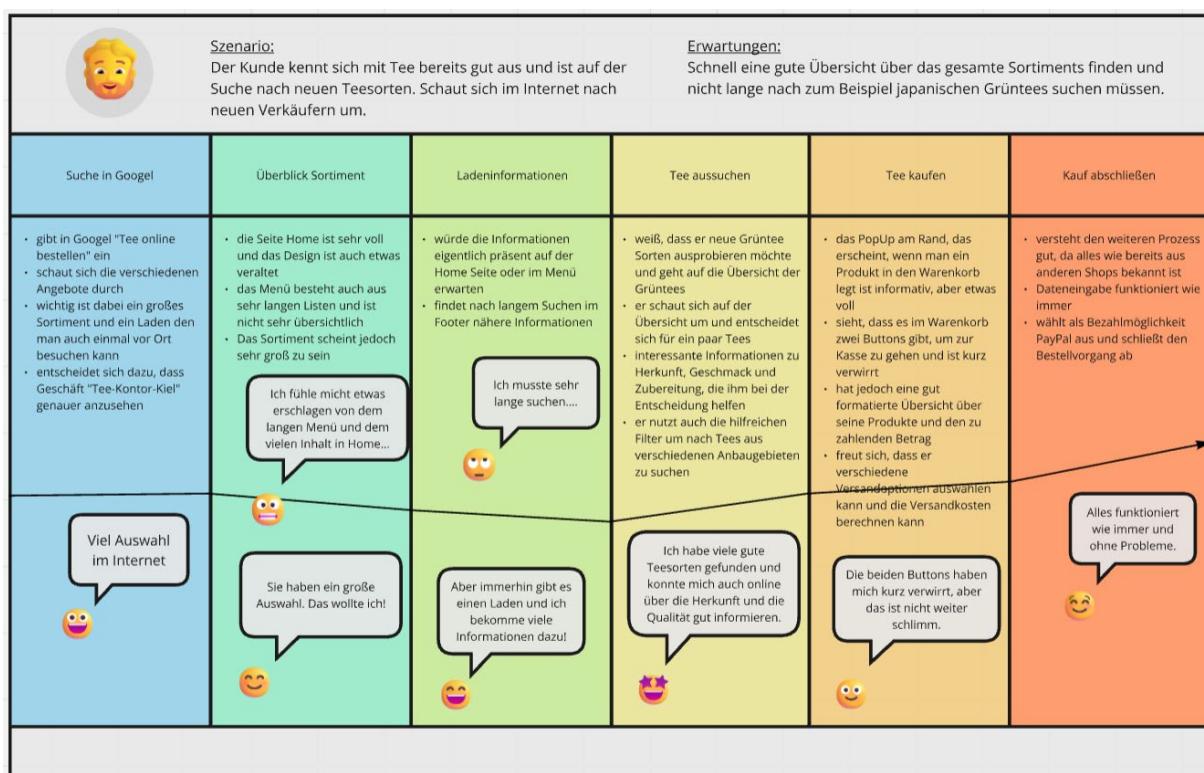


Abbildung 46. Customer Journey Map für Tee Kontor Kiel – Erfahrener Teetrinker

Die zweite Customer Journey Map für Tee Kontor Kiel wurde für den unerfahrenen Teetrinker erstellt, der sich vor allem über Grüntee und dessen Wirkung informieren sowie neue Teesorten ausprobieren möchte. Zunächst wurde ein Szenario entwickelt, indem der Kunde schnell eine gute Übersicht über

das gesamte Sortiment und über die Informationen über die Tees finden möchte. Im nächsten Schritt wurde der Kaufprozess in die einzelnen Schritte unterteilt und in der Map festgehalten. Die wesentlichen Schritte in der Map umfasste die Google-Suche, der Überblick über das Sortiment, das Finden der Ladeninformationen, das Finden der Informationen zu den Tees, das Aussuchen eines Tees und der Abschluss des Kaufes. Bei dieser Map wurde deutlich, dass standardisierte Elemente wie der Warenkorb zwar gut funktionierten, aber unerfahrene Kunden von den langen Menüs und den vielen Inhalten auf einer Seite sehr verunsichert sind und Schwierigkeiten hatten, die gewünschten Informationen schnell zu finden.

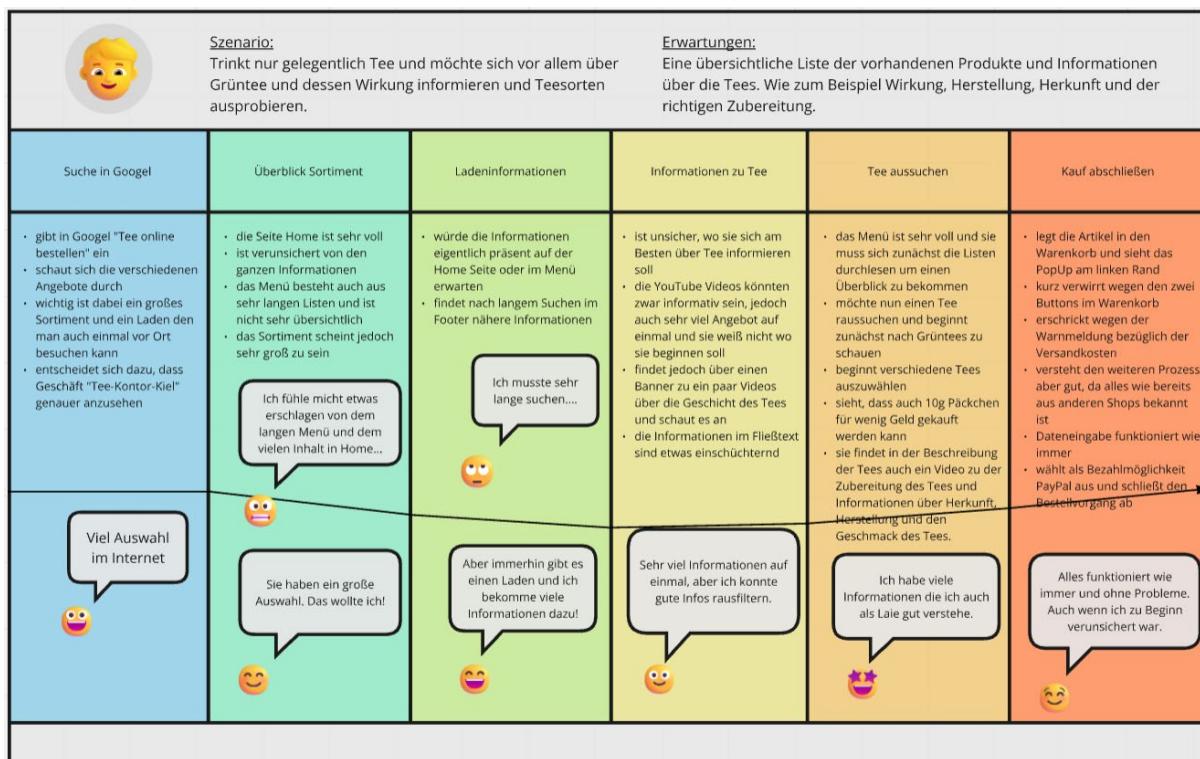


Abbildung 47. Customer Journey Map für Tee Kontor Kiel – Unerfahrener Teetrinker

Anschließend wurde die erste Customer Journey Map für die Website von The English Tearoom und den erfahrenen Teetrinker erstellt, der auf der Suche nach neuen Verkäufern und Teesorten ist. Auch hier wurde ein Szenario entwickelt, in dem der Kunde eine schnelle Übersicht über das gesamte Sortiment erhalten möchte, ohne lange nach bestimmten Tees wie dem japanischen Grüntee suchen zu müssen. Im nächsten Schritt wurde der Kaufprozess in die einzelnen Schritte unterteilt und in der Map festgehalten. Die wesentlichen Schritte waren dabei die Google-Suche, der Überblick über das Sortiment, das Finden der Ladeninformationen, das Aussuchen des Tees, der Kauf des Tees und der Abschluss des Kaufes. Bei dieser Map wurde deutlich, dass sich erfahrene Teetrinker gut in der Auswahl der Tees zurechtfinden, jedoch detailliertere Informationen zu den einzelnen Tees vermissten. Zudem wurde die Website als überladen empfunden, da sie sehr lang ist und mit vielen Effekten arbeitet. Standardisierte Funktionen wie der Warenkorb funktionierten hingegen problemlos.

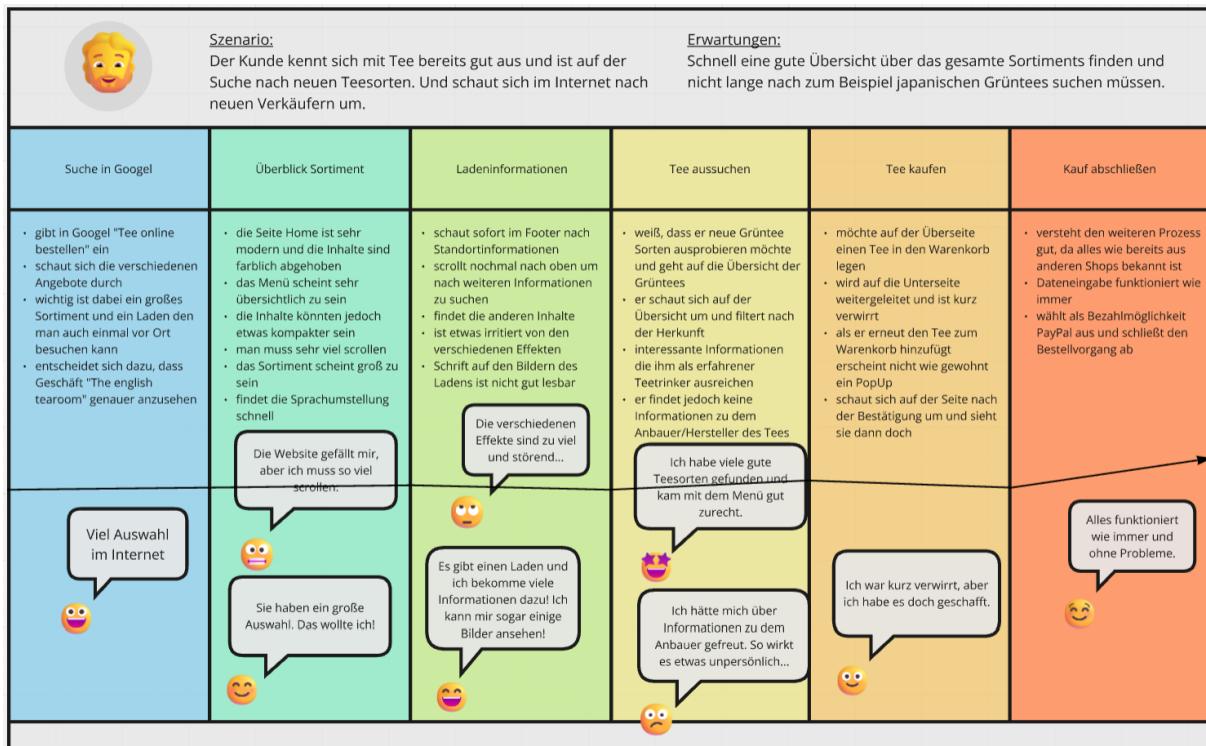


Abbildung 48. Customer Journey Map für The English Tearoom – Erfahrener Teetrinker

Zuletzt wurde die zweite Customer Journey Map für die Website von The English Tearoom und dem unerfahrenen Teetrinker erstellt, der sich über Grüntee und dessen Wirkung informieren sowie neue Teesorten ausprobieren möchte. Auch hier wurde ein Szenario entwickelt, in dem der Kunde eine schnelle Übersicht über das Sortiment sowie gut zugängliche Informationen zu den Tees erwartet. Im nächsten Schritt wurde der Kaufprozess in die einzelnen Schritte unterteilt und in der Map festgehalten. Die wesentlichen Schritte waren dabei die Google-Suche, der Überblick über das Sortiment, das Finden der Ladeninformationen, das Finden der Informationen zu den Tees, das Aussuchen eines Tees und der Abschluss des Kaufes. Bei dieser Map wurde deutlich, dass unerfahrene Teetrinker zunächst durch die visuelle Gestaltung der Website sowie die englische Sprache verunsichert waren. Zudem waren die Ladeninformationen nicht sofort ersichtlich. Die Navigation durch die Teeauswahl hingegen funktionierte reibungslos und die bereitgestellten Informationen zu den einzelnen Tees wurden als ausreichend empfunden.

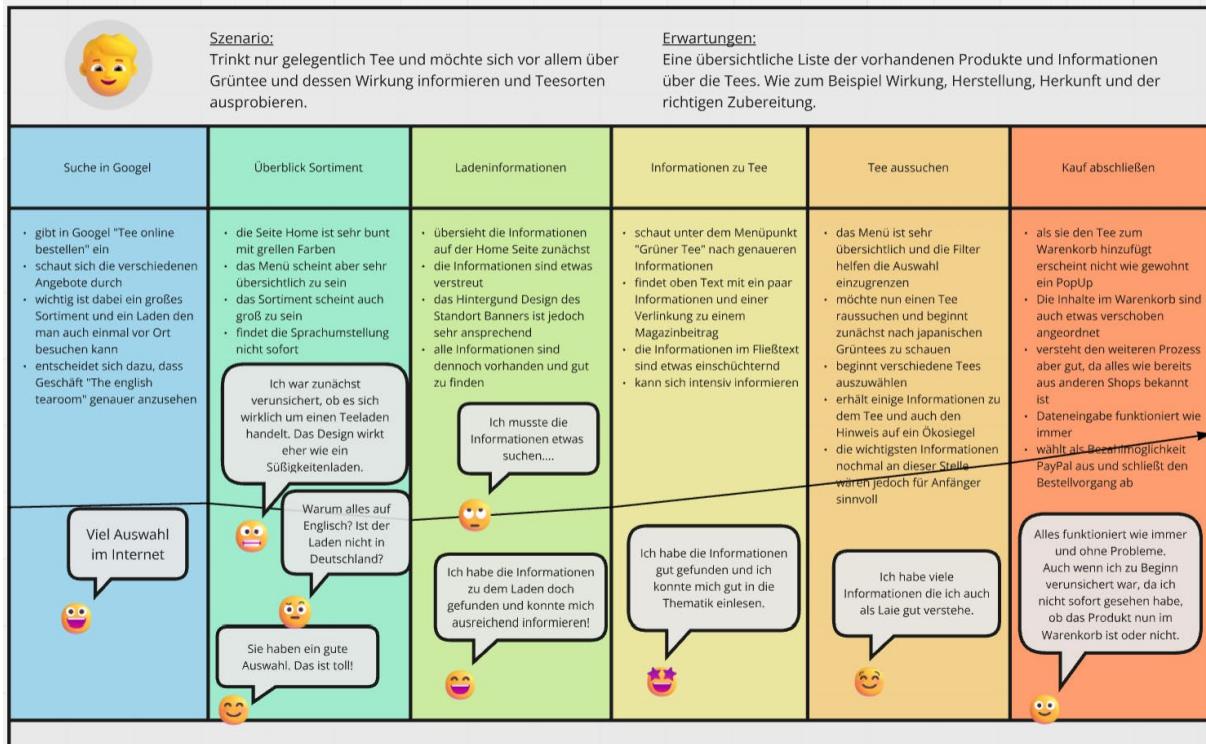


Abbildung 49. Customer Journey Map für The English Tearoom – Unerfahrener Teetrinker

Mithilfe dieser Methode konnten die Nutzerziele in den einzelnen Schritten sowie deren Erfolge und Herausforderungen detailliert erfasst und verständlich dokumentiert werden. Dadurch wurde eine gute Übersicht für die Verbesserungen der eigenen Website geschaffen.

### 3.1.2 Nutzungsanforderungen spezifizieren

Mithilfe der in den folgenden Unterkapiteln beschriebenen Methoden werden die Anforderungen an die neue Applikation festgehalten, wobei der zuvor definierte Nutzungskontext berücksichtigt wird (Burghardt et al., 2011).

#### 3.1.2.1 Personas

Auch für diese Methode wurde zunächst eine Vorlage in Miro erstellt. Anschließend wurde basierend auf den Ergebnissen der Usability Tests und des Feedback Capture Grids die Personas für jede relevante Nutzergruppe ausgefüllt.

Für Kunden mit viel Tee Erfahrung, Kunden mit etwas Tee Erfahrung und Kunden mit so gut wie keiner Tee Erfahrung wurde jeweils eine Persona erstellt. Zunächst wurden für jede Gruppe fiktive, aber dennoch realitätsnahe Namen, Personendaten und Charaktereigenschaften erstellt. Im nächsten Schritt wurde aufbauend auf den vorherigen Methoden ein Standpunkt, Probleme und Pain Points sowie gewünschte Ziele im Kontext festgehalten.

Kunden mit viel Erfahrung bezüglich Tee trinken kennen sich nicht nur mit Tee und dessen Wirkung aus, sondern sind auch meist auf der Suche nach neuen, noch unbekannten Teesorten. Sie möchten sich schnell eine gute Übersicht über das Sortiment verschaffen, ohne lange nach einer speziellen

Teesorte wie japanischen Grüntee suchen zu müssen. Mögliche Probleme und Pain Points könnten dabei ein großes unübersichtliches Menü, fehlerhafte Verlinkungen und ein zu geringes Produktangebot sein sowie zu viele Effekte und zu grelle Farben, die stören und die Website unseriös machen. Zuletzt könnten zu wenig Informationen über die Herkunft und die Anbauer des Tees ein großes Problem darstellen, da erfahrene Teetrinker gerne anhand dieser Kategorien ihre Tees auswählen. Gewünschte Ziele in diesem Kontext könnten daraufhin zum einen ein übersichtliches Menü und gut sortierte Produkte sein. Dies schließt eine sinnvolle Benennung der Menüpunkte und der Kategorien sowie gute und sinnvolle Filtermöglichkeiten mit ein. Zum anderen sollten funktionierende Verlinkungen, ein großes Produktangebot, eine Vergleichsmöglichkeit, ein simples und cleanes Design und keine speziellen und übertriebene Effekte, die ablenken, erzielt werden. Das letzte Ziel sind viele Informationen über die Herkunft und die Anbauer des Tees in visueller Form durch Bilder und Videos sowie in schriftlicher Form.

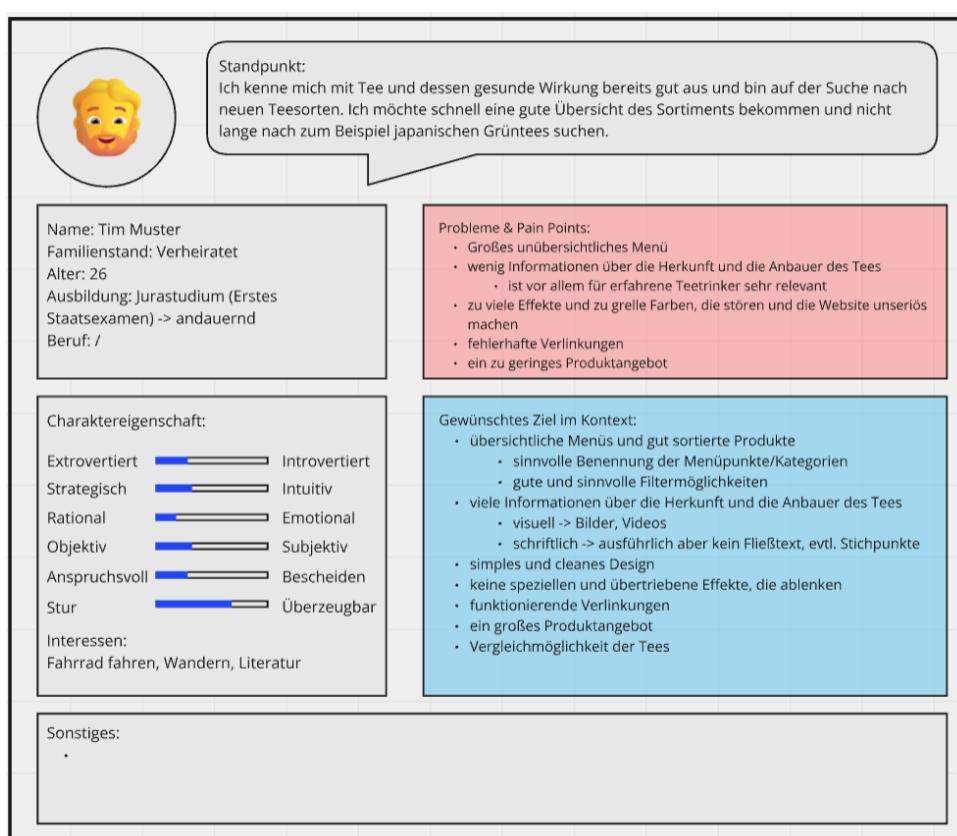


Abbildung 50. Persona des erfahrenen Teetrinkers

Kunden mit etwas Erfahrung bezüglich Teetrinken konsumieren regelmäßig verschiedene Teesorten und interessieren sich neben dem Onlineshop auch für einen physischen Laden, in dem sie sich beraten lassen können. Mögliche Probleme und Pain Points könnten sowohl ein großes unübersichtliches Menü und zu wenig Informationen rund um das Thema Tee im Onlineshop sein, als auch die schwer auffindbaren Informationen über den Laden, da diese meist sehr versteckt oder verstreut sind und selbst mühselig zusammengesucht werden müssen. Weiterhin gibt es zu wenig Informationen über das Verkaufsangebot im Laden. Gewünschte Ziele in diesem Kontext könnten deshalb ein übersichtliches Menü mit guten und sinnvollen Filtermöglichkeiten sowie einer guten Sortierung der Produkte sein. Außerdem

sollten die Informationen über den Laden einfach und präsent auf der Website zu finden sein. Zuletzt sollten viele Informationen rund um das Thema Tee vor allem mit einer Beschreibung der Tees vorhanden sein.

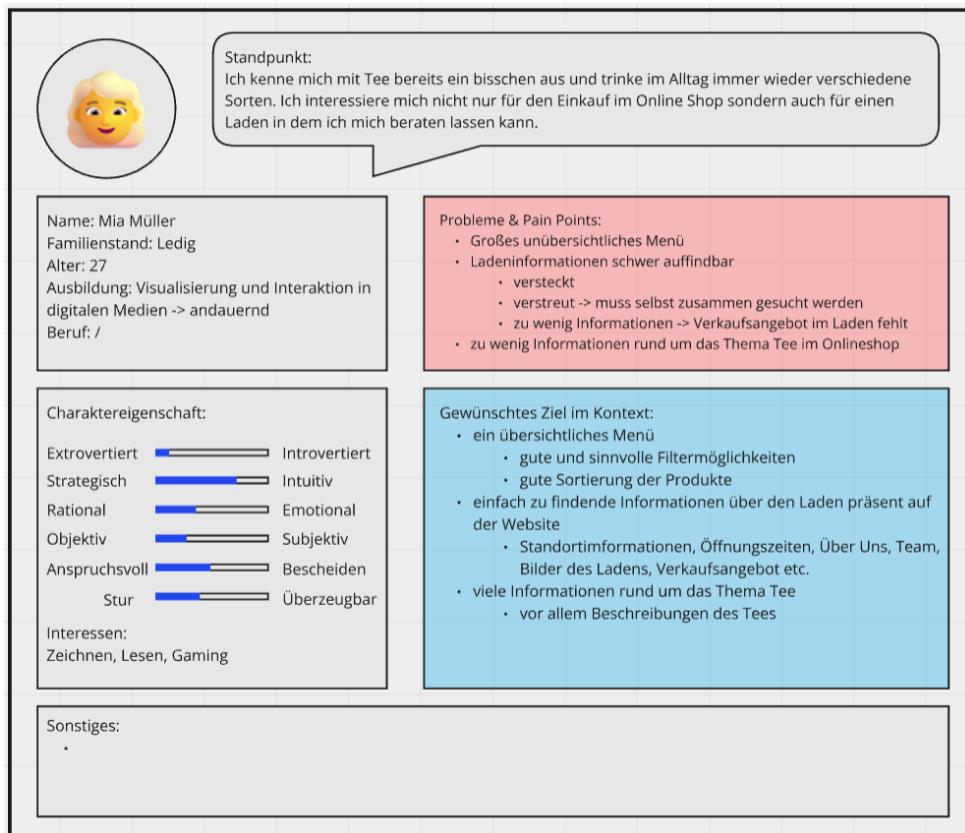


Abbildung 51. Persona des etwas erfahrenen Teetrinkers

Kunden, mit so gut wie keiner Erfahrung bezüglich Tee trinken, konsumieren hauptsächlich Beuteltee aus dem Supermarkt oder gelegentlich losen Grüntee aus dem Bioladen. Sie möchten sich jedoch intensiver mit Tee beschäftigen, mehr über dessen Wirkung, Herstellung und Zubereitung erfahren und neue Sorten ausprobieren. Mögliche Probleme und Pain Points könnten dabei zum einen zu wenig Informationen rund um das Thema Tee und keine Kundenbewertungen sein. Zum anderen können zu viele Informationen auf einmal auch verunsichern. Zuletzt ist eine nicht ansprechende Website ebenfalls ein großes Problem. Dies entsteht durch viel Inhalt auf wenig Platz, ein großes unübersichtliches Menü, zu viele Effekte und unpassende Farben und viele Informationen in Form von Fließtext. Gewünschte Ziele in diesem Kontext könnten daraufhin zum einen viele Informationen rund um das Thema Tee sein, die jedoch kompakt und anschaulich in Form von Stichpunkten und Bildern sowie Videos vermittelt werden sollten. Zum anderen sind Kundenbewertungen gewünscht, um Erfahrung und Empfehlungen anderer einzusehen. Zuletzt sollte die Website ansprechend sein. Das heißt das Design sollte passend sein, nicht zu viele verschiedene Effekte haben, sowie ein übersichtliches Menü und übersichtliche Seiten, die auf unnötige Inhalte verzichten.

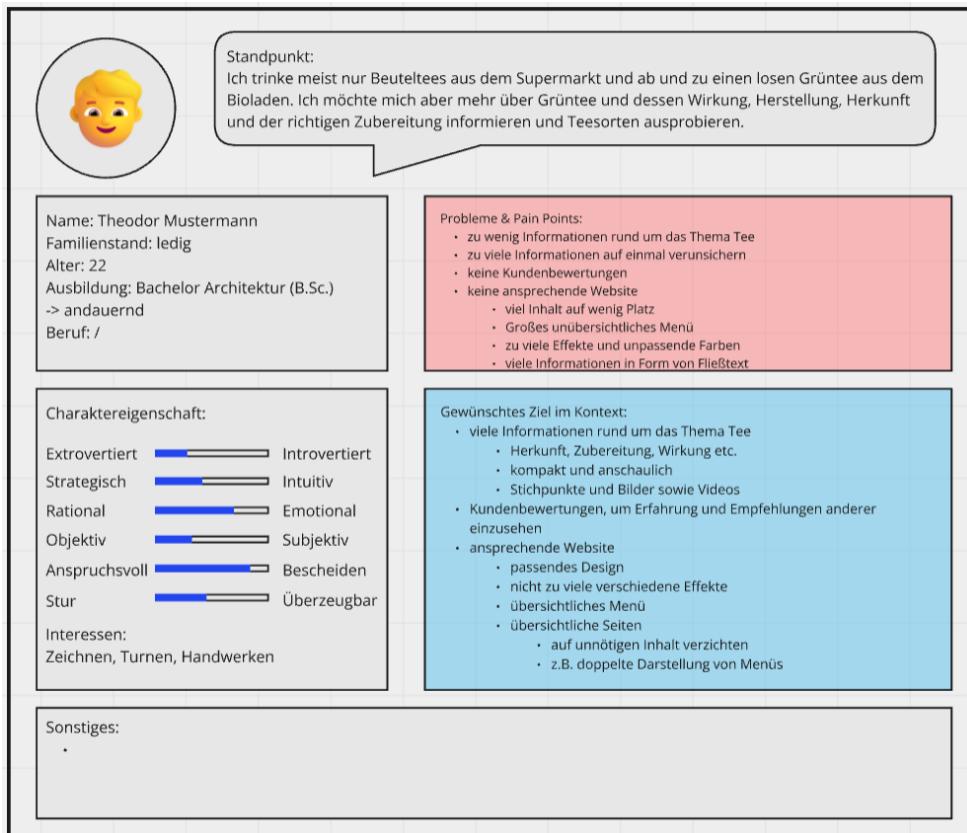


Abbildung 52. Persona des unerfahrenen Teetrinkers

Die Methode half sehr dabei die Erkenntnisse aus vorherigen Methoden übersichtlich festzuhalten. Zudem führte sie zu einem besseren Verständnis der unterschiedlichen Nutzergruppen, ihrer Herausforderungen und Erwartungen.

### 3.1.2.2 User Stories

Zunächst wurde ein passender Leitsatz für die Methode rausgesucht, an dem sich die Bearbeitung orientierte.

Die User Stories wurden auf Grundlage der Beobachtungen aus den Usability Tests, der Umfragen, der Customer Journey Maps und den Personas erstellt. Dabei wurden die einzelnen zuvor ermittelten Probleme durchgegangen und in passende User Stories formuliert.

Die gewonnenen Erkenntnisse deckten sich weitgehend mit den vorherigen Methoden. Alle Nutzergruppen forderten vor allem eine übersichtliche Website, mit einer klar strukturierten Navigation sowie eine verständliche und umfassende Darstellung der relevanten Informationen.

Die Methode half sehr dabei die Anforderungen der Nutzergruppen und deren Hintergründe besser zu verstehen. Zudem bot sie eine klare und einfache Übersicht über die zentralen Bedürfnisse der Nutzer.

Unausgefüllter Leitsatz:

Als **[Nutzer:in/Rolle/Persona]** möchte ich **[ein Ziel/eine Aktion/einen Wunsch]** erreichen, um **[Nutzen]** zu erzielen.

Entwickelte User Stories:

Unerfahrener Teetrinker:

- Als **unerfahrener Teetrinker** möchte ich **viele Informationen rund um Tee auf der Website erreichen, um ein besseres Verständnis für die richtige Zubereitung von Tee und dessen Wirkung zu erzielen.**
- Als **unerfahrener Teetrinker** möchte ich **zunächst kurz und knapp die wichtigsten Informationen präsent angezeigt bekommen**, um **nicht lange suchen und eventuell sogar aufgeben zu müssen.**
- Als **unerfahrener Teetrinker** möchte ich **ein übersichtliches Menü erreichen**, um **die gewünschten Inhalte besser zu finden.**
- Als **unerfahrener Teetrinker** möchte ich **eine übersichtliche Website erreichen**, um **ein besonderes Käuferlebnis zu erzielen.**
- Als **unerfahrener Teetrinker** möchte ich **Kundenbewertungen der Produkte erreichen**, um **Erfahrungen und Empfehlungen von anderen Kunden zu erzielen.**

Etwas erfahrener Teetrinker:

- Als **etwas erfahrener Teetrinker** möchte ich **ein übersichtliches Menü erreichen**, um **die gesuchten Inhalte besser zu finden.**
- Als **etwas erfahrener Teetrinker** möchte ich **schnell eine Übersicht der Ladeninformationen erreichen**, um **den Laden vor Ort besuchen zu können, mich beraten zu lassen und zu sehen, wer hinter dem Onlineshop steht.**
- Als **etwas erfahrener Teetrinker** möchte ich **viele Informationen rund um das Thema Tee erreichen**, um **noch mehr über Tee zu erfahren und zu entscheiden welchen Tee ich am besten kaufen soll.**

Erfahrener Teetrinker:

- Als **erfahrener Teetrinker** möchte ich **viele Informationen über die Herkunft und die Anbauer des Tees erreichen**, um **zu wissen wie, wo und von wem der Tee angebaut wurde, was viel über die Qualität des Tees aussagt.**
- Als **erfahrener Teetrinker** möchte ich **eine seriöse Website ohne viele Effekte und grellen Farben erreichen**, um **ein gutes Käuferlebnis zu erzielen.**
- Als **erfahrener Teetrinker** möchte ich **ein großes Produktangebot erreichen**, um **von vielen verschiedenen Tees die besten rauszusuchen.**
- Als **erfahrener Teetrinker** möchte ich **die Tees vergleichen können**, um **direkt auf der Website die Tees gegenüberzustellen und den besten rauszusuchen.**
- Als **erfahrener Teetrinker** möchte ich **ein übersichtliches Menü mit guten Kategorien und Filtermöglichkeiten erreichen**, um **schnell zu meinen gewünschten Produkten zu finden und ein gutes Käuferlebnis zu erzielen.**

- Als **erfahrener Teetrinker** möchte ich **so wenig Verwirrungen wie möglich auf der Website** erreichen, um **wegen einem schlechten Nutzererlebnis die Seite doch ohne einen Einkauf zu verlassen.**

### 3.1.2.3 How might we...?

Zunächst wurde ein passender Leitsatz für die Methode rausgesucht, an dem sich die Bearbeitung orientieren konnte.

Die Methode How might we...? wurde auf Grundlage der Usability Tests, Umfragen, Customer Journey Maps, Personas und insbesondere der User Stories durchgeführt. Dabei wurden hauptsächlich die User Storys durchgegangen und darauf aufbauend passende Sätze formuliert.

Durch die Methode wurde deutlich, dass der Fokus im weiteren Verlauf insbesondere auf eine allgemein übersichtlichere Website, eine klar strukturierte Navigation sowie eine verständliche und umfassende Informationsdarstellung zum Thema Tee und dessen Herkunft gelegt werden muss. Weiterhin sollten die Forderungen nach Kundenbewertungen, einem großen Produktangebot und einer schnellen und kompakten Übersicht der Ladeninformationen berücksichtigt werden.

Die Methode half sehr dabei einen Grundstein für die Ideenfindung zu legen. Sie bot zudem einen guten Überblick über potenzielle Herausforderungen und Probleme. Diese können nun gut in der Ideenfindung beachtet werden.

Unausgefüllter Leitsatz:

Wie können wir **[ein Ziel/eine Aktion/einen Wunsch/ein Nutzen]** erreichen/verbessern/neu denken?

Entwickelte Fragen:

Unerfahrener Teetrinker:

- Wie können wir **die Vermittlung von Informationen rund um Tee auf der Website** verbessern?
- Wie können wir **präsent, kurz und knapp die Vermittlung der wichtigsten Informationen** erreichen?
- Wie können wir **ein übersichtliches Menü und gut auffindbare Inhalte** erreichen?
- Wie können wir **ein übersichtliches Menü und ein gutes Käuferlebnis** erreichen?
- Wie können wir **Erfahrungen und Empfehlungen in Form von Kundenbewertungen** erreichen?

Etwas erfahrener Teetrinker:

- Wie können wir **ein übersichtliches Menü** erreichen/verbessern?
- Wie können wir **eine schnelle und kompakte Übersicht der Ladeninformationen mit allen wichtigen Details** erreichen/verbessern?
- Wie können wir eine **Informationsvermittlung rund um das Thema Tee verständlich und übersichtlich** erreichen/verbessern?

Erfahrener Teetrinker:

- Wie können wir **die Vermittlung vieler Informationen über die Herkunft und die Anbauer des Tees** erreichen/verbessern?
- Wie können wir **eine seriöse Website** erreichen/verbessern/neu denken?
- Wie können wir **ein großes Produktangebot** erreichen?
- Wie können wir **eine Vergleichsmöglichkeit der Tees** erreichen/neu denken?
- Wie können wir **ein übersichtliches Menü mit guten Kategorien und Filtermöglichkeiten** erreichen/verbessern?
- Wie können wir **so wenig Verwirrungen wie möglich auf der Website** erreichen?

### 3.1.3 Lösungsentwicklung auf Basis der Nutzungsanforderungen

Mithilfe der in den folgenden Unterkapiteln beschriebenen Methoden werden Gestaltungslösungen auf Basis der zuvor definierten Nutzungsanforderungen sowie den Gesetzen und Prinzipien des Webdesigns entwickelt. Die Lösungen werden anschließend in Form von Prototypen konkretisiert (Burghardt et al., 2011).

#### 3.1.3.1 Brainstorming

Zunächst wurde eine passende Vorlage für die Methode in Miro vorbereitet. Anschließend wurde auf Basis der Personas, Customer Journey Maps, User Stories und der How might we...? Methode erste Ideen gesammelt und unsortiert notiert. Um eine bessere Übersicht zu erhalten, wurden die Ideen anschließend den beiden Oberbegriffen Gestaltung und Inhalt zugeordnet.

Im Folgenden werden erste Ideen vorgestellt, die eine gute Basis für die mögliche Anwendung geben werden. Unter dem Oberbegriff Inhalt wurden Ideen, wie eine Vergleichsmöglichkeit der Produkte entwickelt. Dies könnte in einer Tabelle mit individuellen Spalten erfolgen, in der die wichtigsten Daten wie Herkunft und Geschmack verglichen werden können. Kategorien sind zudem ergänz- und löscharbar. Weiterhin könnte man Probepackungen der Tees anbieten, damit sich Kunden durch das Angebot probieren können und nicht von hohen Preisen abgeschreckt werden. Eine extra Seite „Alles rund um Tee“ soll dem Kunden eine Information direkt auf der Website ermöglichen. Dafür soll zum einen ein extra Menüpunkt unter Tee eingerichtet werden sowie eine Verlinkung auf Home. In dem Bezug sollte aber auch darauf geachtet werden, kurze Seiten mit kompaktem Inhalt zu erstellen. Auch für die Informationen über den Laden sollte zusätzlich zu den Informationen auf der Home Seite eine extra Seite angelegt werden mit detaillierteren Inhalten sowie mit Bildern und Videos. Es sollte zudem Informationen über das Verkaufsangebot im Ladengeschäft geben. Wichtige Elemente, die von jeder Seite zugänglich sein sollten, wie zum Beispiel die Sprachumstellung, der Warenkorb und die Suche sollte sehr präsent, wie zum Beispiel im Header über dem Menü, platziert werden. Zudem sollte es an passenden Stellen Verlinkungen zu YouTube oder Google Maps geben. Jedoch sollte vor einer Weiterleitung eine Abfrage erfolgen, bei der der Nutzer einer Weiterleitung erneut zustimmen muss, um ein Versehen und Frustrationen vorzubeugen. Weiterhin sollten wichtige Daten bei jedem Tee angezeigt werden, wie der Herkunftsor sowie der Anbauer, eine kurze Hintergrundgeschichte, die Zubereitung, der Geschmack, die Wirkung und allgemeine Informationen zu der Teeart. Dabei sollte es nicht nur Fließtext, sondern auch

Stichpunkte geben, die durch Bilder oder Videos unterstützt werden. Zusätzlich sollte das Feedback von anderen Kunden präsent gezeigt werden sowie Empfehlungen beziehungsweise „andere Kunden kauften auch“ oder „haben sich ebenfalls angesehen“ sollte mit aufgenommen werden. Dabei sollten die wichtigsten Daten bereits in der Übersicht der Produkte angezeigt werden und nicht nur der Name sowie der Preis des jeweiligen Produktes. Zuletzt könnte man Gutscheine anbieten. Die Kunden können den Betrag zum einen in fünf Euro Schritten auswählen oder durch eine freie beliebige Eingabe des Geldbetrags.

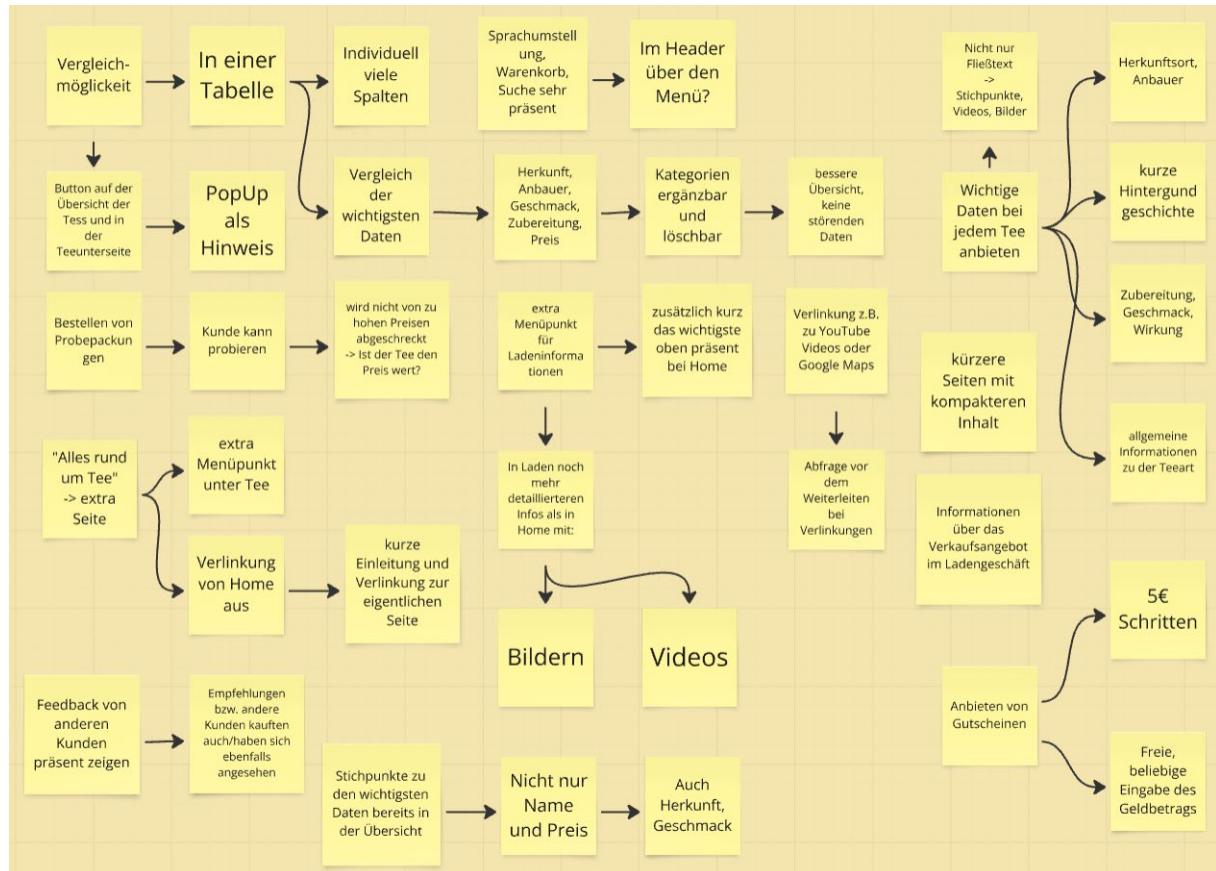


Abbildung 53. Ideen aus dem Brainstorming unter dem Oberbegriff Inhalt

Unter dem Oberbegriff Gestaltung wurden Ideen wie ein simples Design entwickelt, dass aus dezenten nicht zu vielen verschiedenen Farben besteht, die nicht zu grell sind. Weiterhin sollte auf viele verschiedene Gestaltungselemente verzichtet werden. Passend dazu wurde die Idee entwickelt, eine leichte farbliche Unterscheidung je nach Kategorie vorzunehmen. Da dies jedoch im Widerspruch zu der obigen Idee steht, muss eine der beiden Ideen verworfen werden. Auch auf unnötige Elemente und Effekte sollte verzichtet werden, das heißt keine doppelten Buttons mit der gleichen Funktion auf einer Seite und nicht zu viele Effekte, die den Nutzer verwirren. Auch auf scrollbare Elemente in der Seite sollte verzichtet werden. Weiterhin sollte das Menü übersichtlich gestaltet werden, indem zunächst nur Oberkategorien angezeigt werden, die durch weitere aufklappbare Menülisten oder durch Filter in den Unterseiten weiter eingegrenzt werden. Weitere Ideen wie das Essensangebot vor Ort in Form einer typischen Karte zu gestalten und das ermöglichen eines deutlichen Feedbacks in Form von Pop-ups oder präsente Meldungen wurden ebenfalls entwickelt. Zudem sollte auf genug Abstände zwischen den

Elementen geachtet werden, um eine Abgrenzung zu gewährleisten. Um die Tees ansprechend zu präsentieren, sollte auf ansprechende Bilder beziehungsweise auf Bilder des Tees mit Verpackung geachtet werden. Zuletzt könnten Produkte im Warenkorb angeboten beziehungsweise verlinkt werden, die der Kunde noch kaufen könnte, um einen kostenlosen Versand zu erreichen.

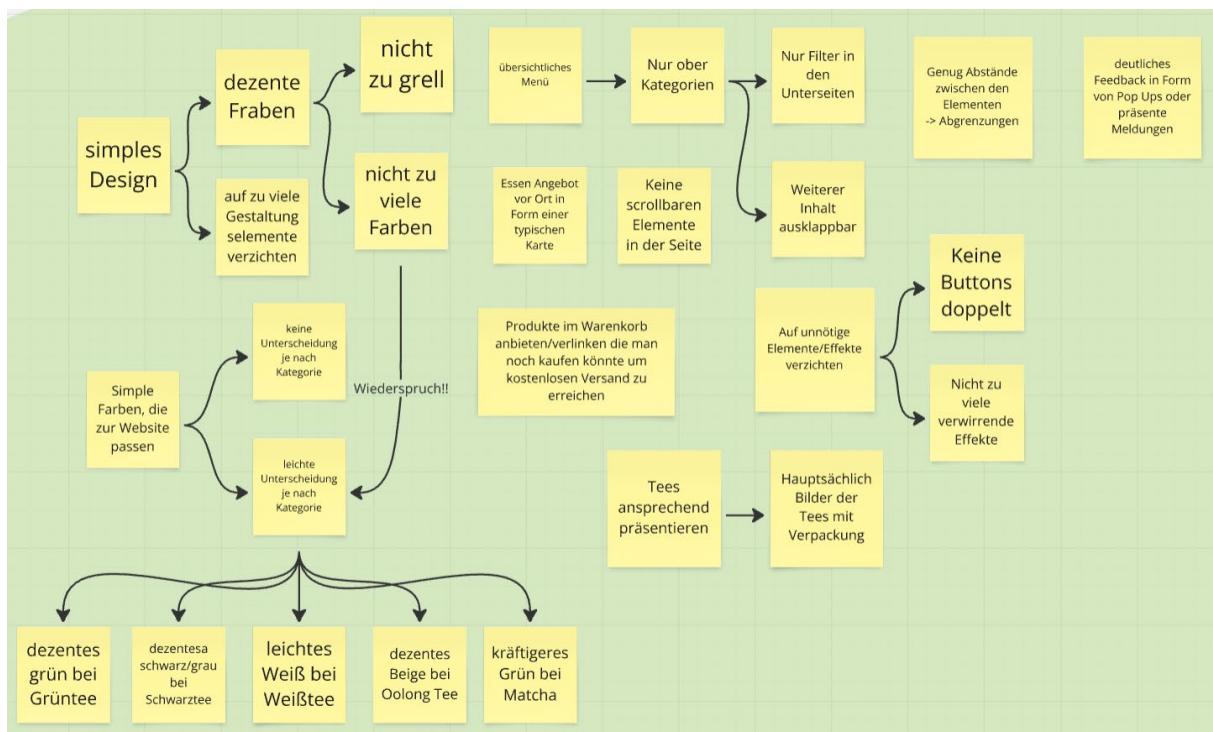


Abbildung 54. Ideen aus dem Brainstorming unter dem Oberbegriff Gestaltung

Die Methode ermöglichte eine vielseitige und kreative Denkweise. Dadurch konnten zahlreiche neue Ansätze sowohl für den inhaltlichen Aufbau als auch für die visuelle Gestaltung der Anwendung identifiziert werden.

### **3.1.3.2 Hybrides Card Sorting**

Zunächst wurde eine passende Vorlage für die Methode in Miro vorbereitet. Anschließend wurden die erforderlichen Inhalte der neuen Anwendung basierend auf den vorherigen Methoden sowie der Auftragsstellung des fiktiven Kunden gesammelt. Diese Inhalte wurden nach ihrer thematischen Zugehörigkeit sortiert und in sinnvolle Oberbegriffe gegliedert. Daraus entstand eine erste Informationsstruktur für die neue Anwendung.

Es ergaben sich schlussendlich folgende Oberbegriffe: Home, Tee, Matcha, Zubehör, Köstlichkeiten, Laden, Gutscheine und Header allgemein. Home enthält Informationen zum Ladengeschäft, allgemeine Infos rund um Tee, Bestseller und Neuheiten an Tees sowie Zubehör und eine Kontaktmöglichkeit.

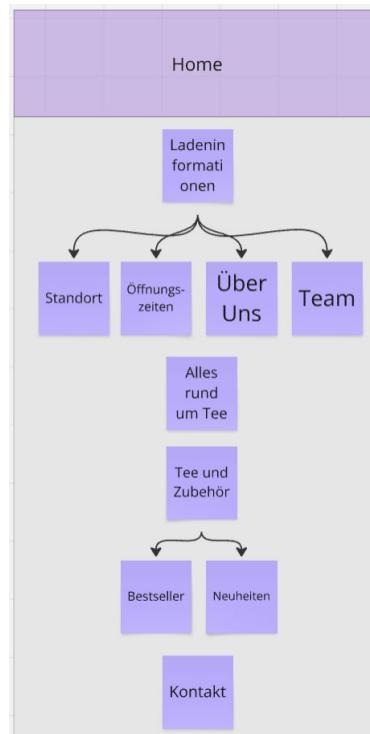


Abbildung 55. Hybrides Card Sorting – Oberbegriff Home

Der Oberbegriff Tee umfasst die Kategorien Alles rund um Tee, Grüner Tee, Schwarzer Tee, Weißer Tee, Oolong Tee, Früchtetee und Kräutertee mit den entsprechenden Filtern und Unterkategorien.



Abbildung 56. Hybrides Card Sorting – Oberbegriff Tee

Dem Oberbegriff Matcha wurde Alles über Matcha, Matcha-Pulver, Matcha-Zubehör und Matcha-Sets, ebenfalls mit passenden Filtern und Unterkategorien zugeteilt.

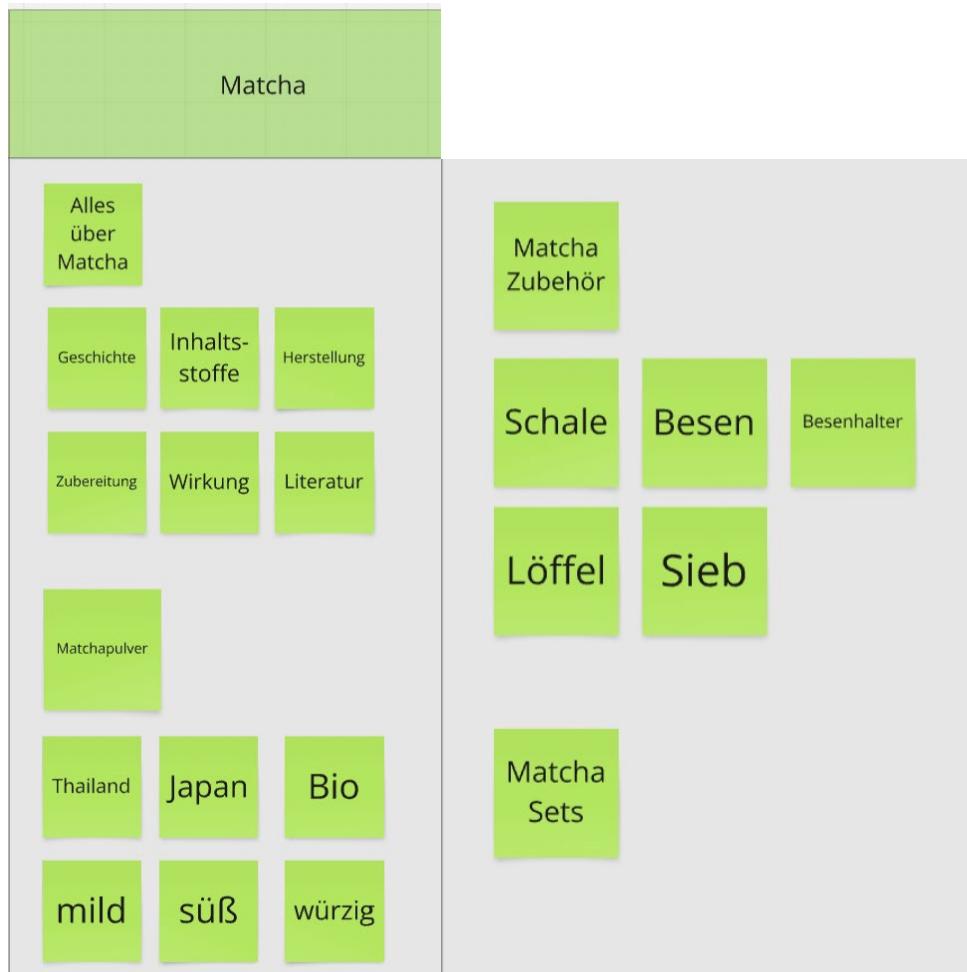


Abbildung 57. Hybrides Card Sorting – Oberbegriff Matcha

Für den Oberbegriff Zubehör entwickelten sich die Inhalte Teedosen, Teeflaschen, Wasserkocher und Teegeschirr.



Abbildung 58. Hybrides Card Sorting – Oberbegriff Zubehör

Für die Köstlichkeiten ergab sich eine Aufteilung in süß und herhaft zu denen anschließend passende Produkte rausgesucht wurden.

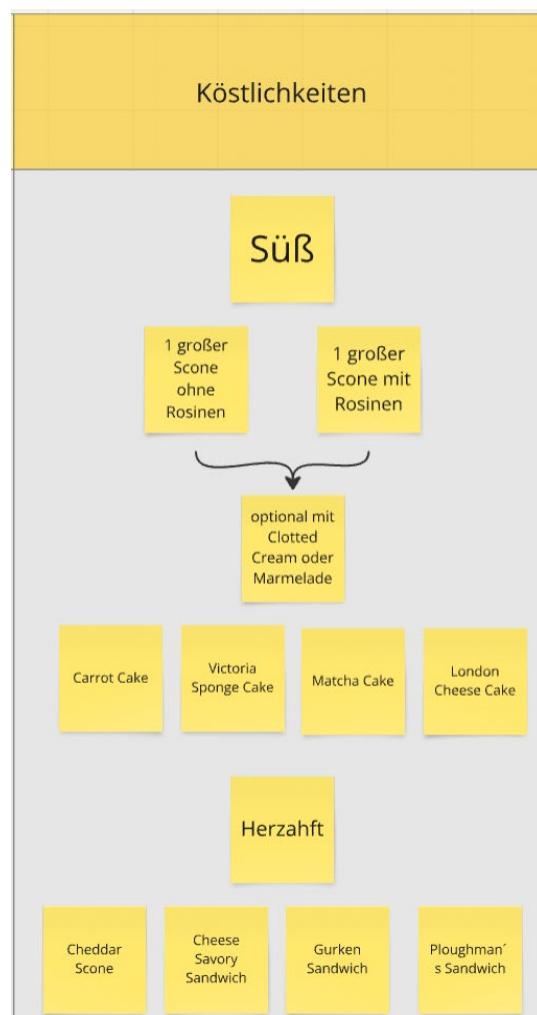


Abbildung 59. Hybrides Card Sorting – Oberbegriff Köstlichkeiten

Unter dem Oberbegriff Laden sammelten sich die Inhalte Standort, Öffnungszeiten, Über Uns, Team und Kontakt.



Abbildung 60. Hybrides Card Sorting – Oberbegriff Laden

Für die Gutschein wurde eine Untergrenze von fünf Euro und ein Maximalbetrag von hundertfünfzig Euro festgelegt.



Abbildung 61. Hybrides Card Sorting – Oberbegriff Gutscheine

Der Oberbegriff Header Allgemein enthält zentrale Navigationselemente wie Suche, Logo, Warenkorb, Sprachumstellung, Merkliste und Vergleichsliste.

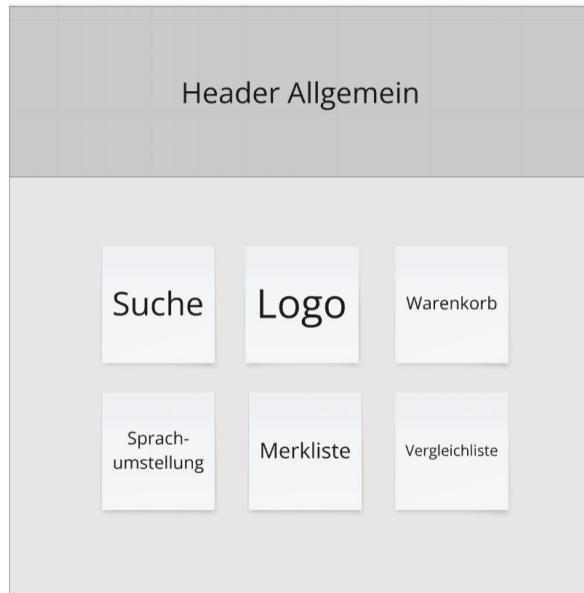


Abbildung 62. Hybrides Card Sorting – Oberbegriff Header Allgemein

Durch das Hybride Card Sorting konnte zunächst eine umfassende Übersicht über die notwendigen Inhalte erstellt werden. Darauf aufbauend wurde eine logische und benutzerfreundliche Informationsstruktur entwickelt, die eine solide Basis für die kommenden Prototypen bietet.

### 3.1.3.3 Future-State Journey Map

Zunächst wurde eine passende Vorlage für die Methode in Miro vorbereitet. Anschließend wurden auf Basis des Brainstormings und des Card Sortings zwei Future-State Journey Maps erstellt. Eine Map wurde für die Nutzergruppe der erfahrenen Teetrinker entwickelt, während die zweite Map die Nutzergruppen der etwas erfahrenen und unerfahrenen Teetrinker gemeinsam abbildet. Da sich diese beiden Gruppen stark ähneln, gab es keine relevanten Unterschiede, die eine separate Map erforderlich gemacht hätten. Für die Inhalte des Szenarios und der Erwartungen der Future-State Journey Maps wurden die bereits erarbeiteten Inhalte aus den Customer Journey Maps erneut verwendet. Diese waren weiterhin sinnvoll und ermöglichten eine präzisere Verdeutlichung des Verbesserungspotenzials.

Zunächst wurde eine Map für den erfahrenen Teetrinker entwickelt. Dieser ist auf der Suche nach neuen Teesorten und Anbietern und möchte sich schnell einen Überblick über das Sortiment verschaffen, ohne lange nach spezifischen Tees wie dem japanischen Grüntee suchen zu müssen. Im nächsten Schritt wurde die Dokumentation in die einzelnen Schritte unterteilt und in der Map festgehalten. Die wesentlichen Schritte waren dabei die Google-Suche, der Überblick über das Sortiment, das Auffinden der Ladeninformationen, das Aussuchen des Tees, der Kauf des Tees und der Abschluss des Kaufes. Die meisten Unterschiede, mit dem größten Verbesserungspotenzial liegen in der Übersichtlichkeit der Website, dem schnellen Auffinden relevanter Informationen über Teesorten und das Ladengeschäft sowie einem verständlichen und effizienten Kaufprozess. Besonders vorteilhaft für diese Nutzergruppe ist die Vergleichsfunktion, die es ermöglicht, verschiedene Tees auf einen Blick miteinander zu vergleichen.

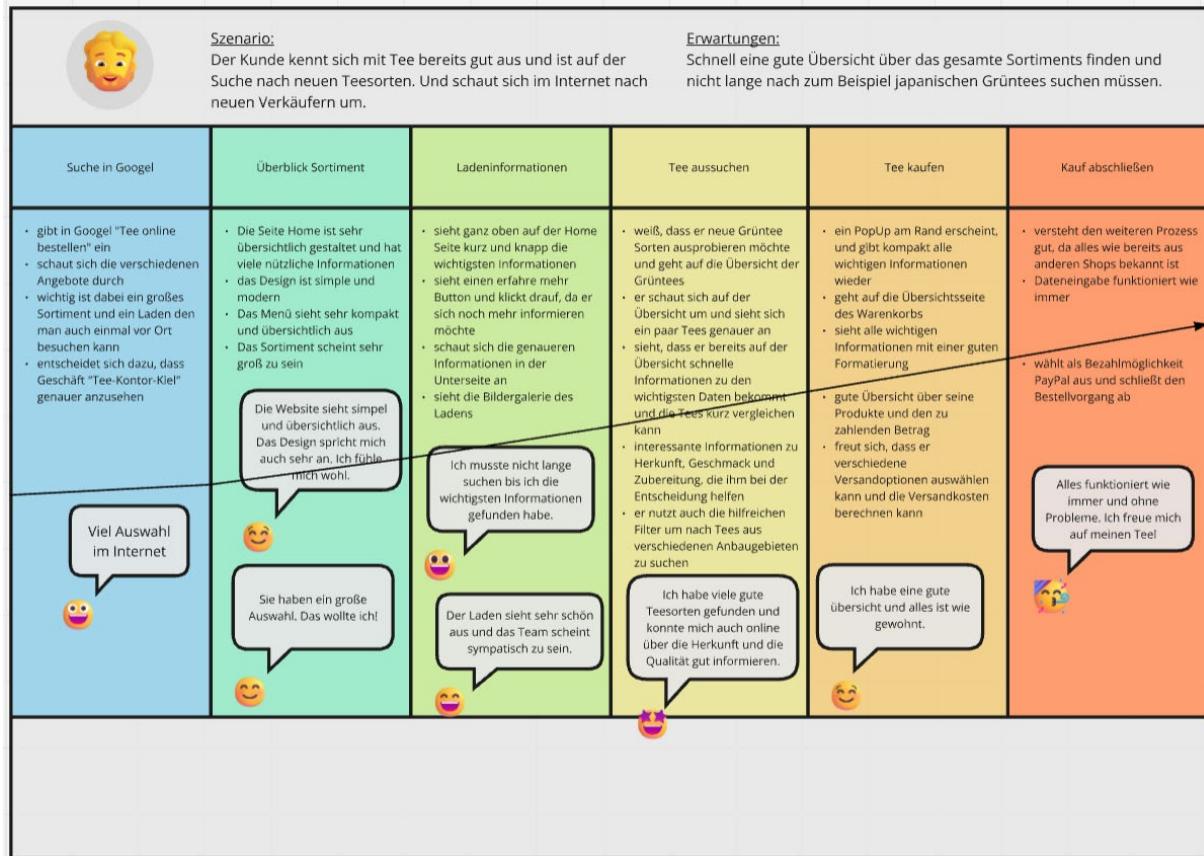


Abbildung 63. Future Journey Map – Erfahrener Teetrinker

Die zweite Future-State Journey Map fokussiert sich auf den unerfahrenen Teetrinker, der sich über Grüntee und dessen Wirkung informieren sowie neue Teesorten ausprobieren möchte. Zunächst wurde ein Szenario entwickelt, indem der Kunde schnell eine gute Übersicht über das gesamte Sortiment und über die Informationen der Tees finden möchte. Im nächsten Schritt wurde die Dokumentation in die einzelnen Schritte unterteilt und in der Map festgehalten. Die wesentlichen Schritte waren dabei die Google-Suche, der Überblick über das Sortiment, das Auffinden der Ladeninformationen, das Finden der Informationen zu den Tees, das Aussuchen eines Tees und der Abschluss des Kaufes. Auch hier zeigten sich die größten Unterschiede in der Struktur und Übersichtlichkeit der Website, der leichten Auffindbarkeit von Informationen und einem intuitiven Kaufprozess. Besonders vorteilhaft für diese Nutzergruppe ist die Kategorie „Alles über Tee“, die eine kompakte und übersichtliche Sammlung wichtiger Informationen bietet, ohne dass der Nutzer lange suchen muss.

 <p><b>Szenario:</b> Trinkt nur gelegentlich Tee und möchte sich vor allem über Grüntee und dessen Wirkung informieren und Teesorten ausprobieren.</p> <p><b>Erwartungen:</b> Eine übersichtliche Liste der vorhandenen Produkte und Informationen über die Tees. Wie zum Beispiel Wirkung, Herstellung, Herkunft und der richtigen Zubereitung.</p>					
Suche in Google	Überblick Sortiment	Ladeninformationen	Informationen zu Tee	Tee aussuchen	Kauf abschließen
<ul style="list-style-type: none"> <li>gibt in Google "Tee online bestellen" ein</li> <li>schaudt sich die verschiedenen Angebote durch</li> <li>wichtig ist dabei ein großes Sortiment und ein Laden den man auch einmal vor Ort besuchen kann</li> <li>entscheidet sich dazu, dass Geschäft "Tee-Kontor-Kiel" genauer anzusehen</li> </ul>	<ul style="list-style-type: none"> <li>die Seite Home ist sehr übersichtlich gestaltet und hat viele nützliche Informationen</li> <li>das Design ist simple und modern</li> <li>das Menü sieht sehr kompakt und übersichtlich aus</li> <li>das Sortiment scheint sehr groß zu sein</li> </ul> <p><i>(Die Website sieht simpel und übersichtlich aus. Das Design spricht mich auch sehr an. Ich fühle mich wohl.)</i></p>	<ul style="list-style-type: none"> <li>sieht ganz oben auf der Home Seite kurz und knapp die wichtigsten Informationen</li> <li>sieht einen "Erfahre mehr" Button und klickt drauf, da sie sich noch mehr informieren möchte</li> <li>schaudt sich die genaueren Informationen in der Unterseite an</li> <li>sieht die Bildergalerie des Ladens</li> </ul>	<ul style="list-style-type: none"> <li>sieht unter dem Menüpunkt Tee eine Kategorie "Alles über Tee"</li> <li>bekommt zu jeder Teesorte kurze und interessante Informationen</li> <li>Empfehlungen zu Websites, Büchern oder Videos um sich noch genauer zu informieren</li> <li>Informationen werden in Form von Stichpunkten und Fließtext sowie mit Bildern und Videos als Unterstützung vermittelt</li> </ul> <p><i>(Ich bekomme eine gute Übersicht über die wichtigsten Informationen und habe sogar Empfehlungen für das Aufinden von mehr Informationen bekommen!)</i></p>	<ul style="list-style-type: none"> <li>das Menü ist sehr übersichtlich und sie findet sich sehr schnell zurecht</li> <li>möchte nun einen Tee raus suchen und beginnt zunächst nach Grüntees zu schauen</li> <li>beginnt verschiedene Tees auszuwählen</li> <li>sieht, dass auch 10g Packchen für wenig Geld gekauft werden kann</li> <li>sieht die Möglichkeit die Tees zu vergleichen und vergleicht 3 verschiedene Tees</li> <li>kann sich so für 2 entscheiden</li> </ul> <p><i>(Die Auswahl war durch die übersichtliche Webseite und die Vergleichsmöglichkeit schnell und gut möglich.)</i></p>	<ul style="list-style-type: none"> <li>ein PopUp am Rand erscheint, und gibt kompakt alle wichtigen Informationen</li> <li>geht auf die Übersichtsseite des Warenkorbs</li> <li>sieht alle wichtigen Informationen mit einer guten Formattierung</li> <li>gute Übersicht über seine Produkte und den zu erreichenden Betrag</li> <li>versteht den weiteren Prozess aber gut, da alles wie bereits aus anderen Shops bekannt ist</li> <li>Dateneingabe funktioniert wie immer</li> <li>wählt als Bezahlmöglichkeit PayPal aus und schließt den Bestellvorgang ab</li> </ul> <p><i>(Alles funktioniert wie immer und ohne Probleme. Das war einfach und ging schnell.)</i></p>
<i>Viel Auswahl im Internet</i>	<i>😊</i>	<i>😊</i>	<i>😊</i>	<i>😊</i>	<i>😊</i>

Abbildung 64. Future Journey Map – Unerfahrener Teetrinker

Diese Methode hat sehr dabei geholfen, einen konkreten Ablauf für die verbesserte Website zu entwickeln. Die Future-State Journey Maps haben gezeigt, dass die erarbeiteten Ideen aus dem Brainstorming sinnvoll sind und sich in der neuen Anwendung bewähren werden. Sie verdeutlichen das Verbesserungspotenzial und bieten eine klare Grundlage für die weitere Umsetzung.

### 3.1.3.4 Informationsarchitektur

Die Informationsarchitektur wurde in Miro basierend auf den vorherigen Methoden, insbesondere auf Basis des Card Sortings erstellt. Dabei wurden der Header, der Footer-Inhalt, die einzelnen Seiten, deren Inhalte und interaktive Elemente wie Filter und Suchfunktionen strukturiert dargestellt.

Dabei wurde das Card Sorting Schritt für Schritt durchgegangen und verbessert, wodurch sich nach und nach eine gute Struktur ergab. Zusätzlich wurden die Erkenntnisse aus der Future-State Journey Map, dem Brainstorming sowie den Nutzerbedürfnissen aus den Methoden Persona, User Stories und How might we...? berücksichtigt. Besondere Aufmerksamkeit galt der sinnvollen Kategorisierung des Produktangebotes, der Header-Inhalte und der Strukturierung der Startseite.

Diese Methode erwies sich als äußerst hilfreich, um die Inhaltsstruktur der Website visuell und übersichtlich darzustellen und gezielt für eine benutzerfreundliche Anwendung zu verbessern.

### 3.1.3.5 Wireframing

Die Methode wurde auf Basis des Brainstormings durchgeführt, um die ersten Ideen visuell darzustellen.

Die Wireframes wurden dabei von Hand in Good Notes mit einem Apple Pencil skizziert.

Durch die Erstellung der Wireframes konnten Ideen konkretisiert, weiterentwickelt oder verworfen werden. Die Idee das Essensangebot in Form einer Karte darzustellen, wurde verworfen, da die Bilder nur in kleiner Größe abgebildet werden können, dieser Entwurf der Idee des simplen Designs widersprechen würde, die Seite sehr unordentlich wirkte und kein einheitliches Design mit den anderen Teeseiten herrschte. Obwohl eine Abgrenzung zwischen dem Essensangebot und dem Onlineshop theoretisch von Vorteil sein könnte, überwogen die negativen Aspekte dieser Umsetzung.

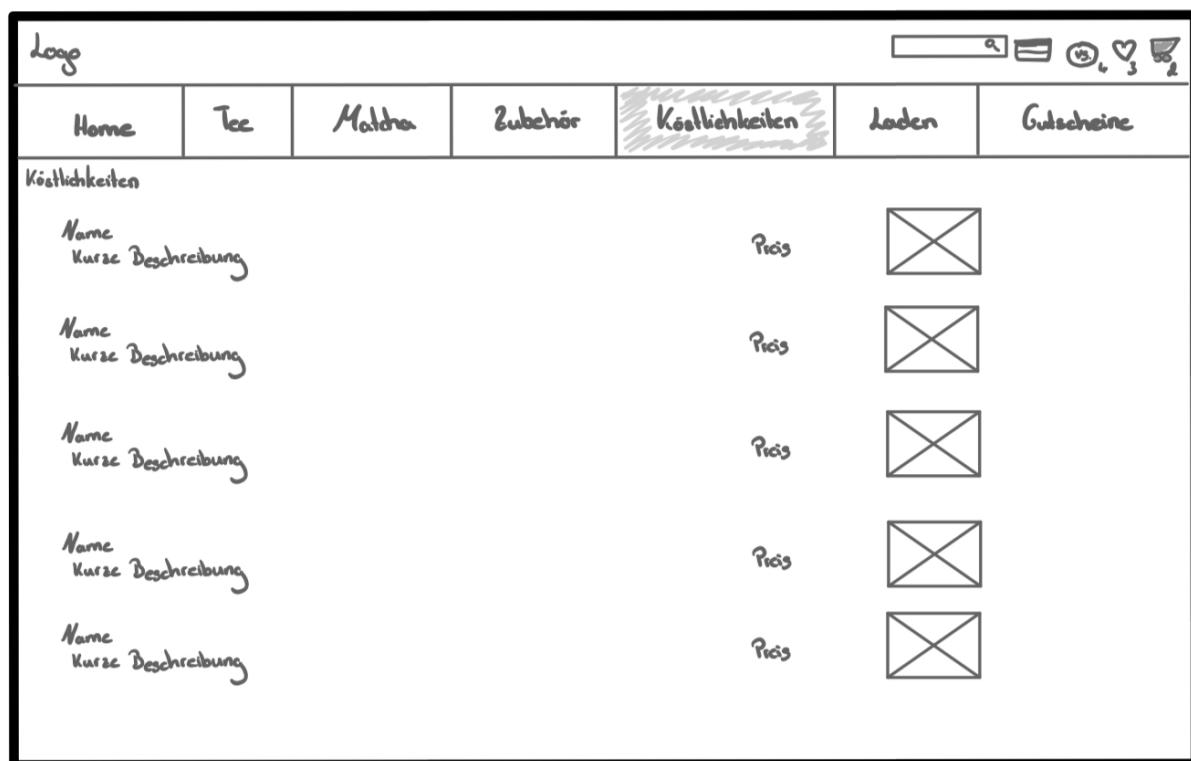


Abbildung 65. Erster Entwurf der Seite „Köstlichkeiten“

Daraufhin wurde ein zweiter Entwurf der Seite „Köstlichkeiten“ entworfen, der sich optisch am Onlineshop orientiert. Jedoch wurden keine Filtermöglichkeiten integriert, da diese hier nicht benötigt werden. Zudem ist es dem Nutzer nicht möglich die Produkte in den Warenkorb zu legen, da die Produkte ausschließlich vor Ort im Laden erhältlich sind. Um das Verständnis der Nutzer zu verbessern, ist ein deutlicher Hinweis diesbezüglich wichtig.

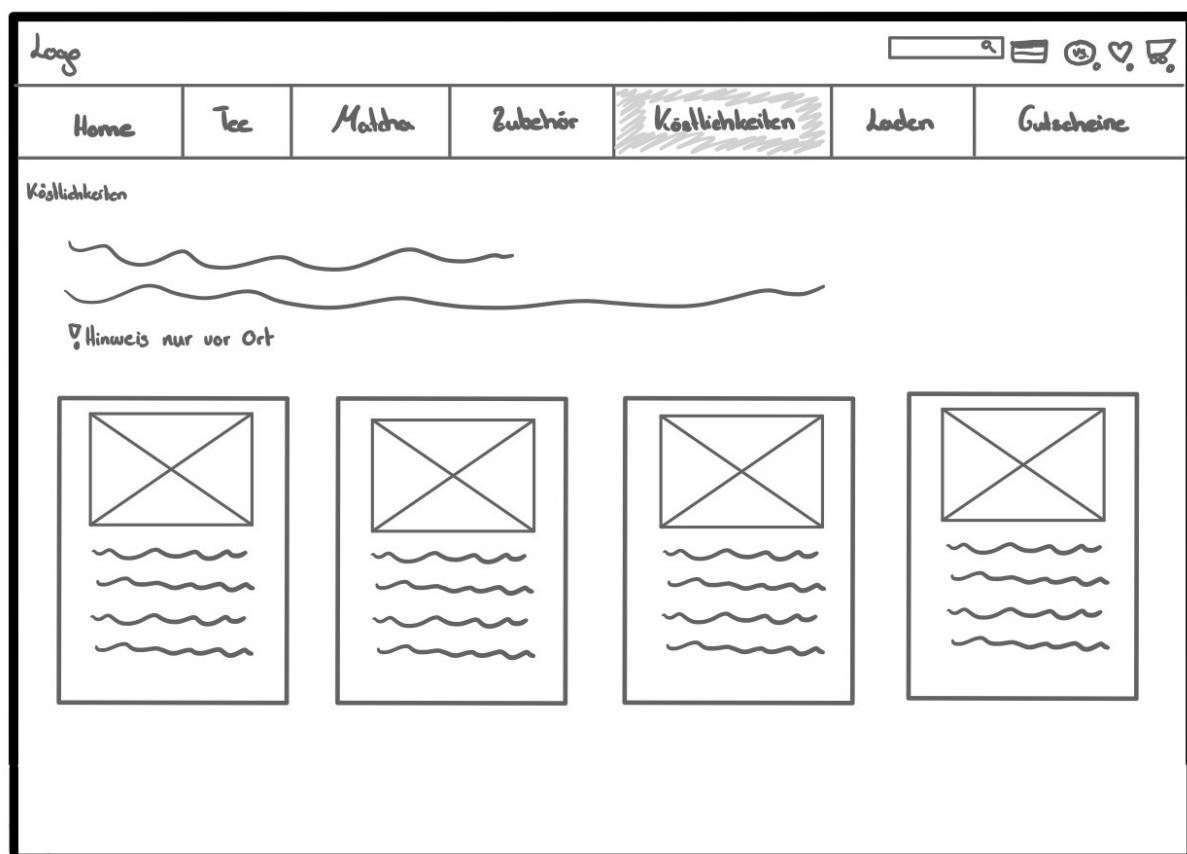


Abbildung 66. Zweiter Entwurf der Seite „Köstlichkeiten“

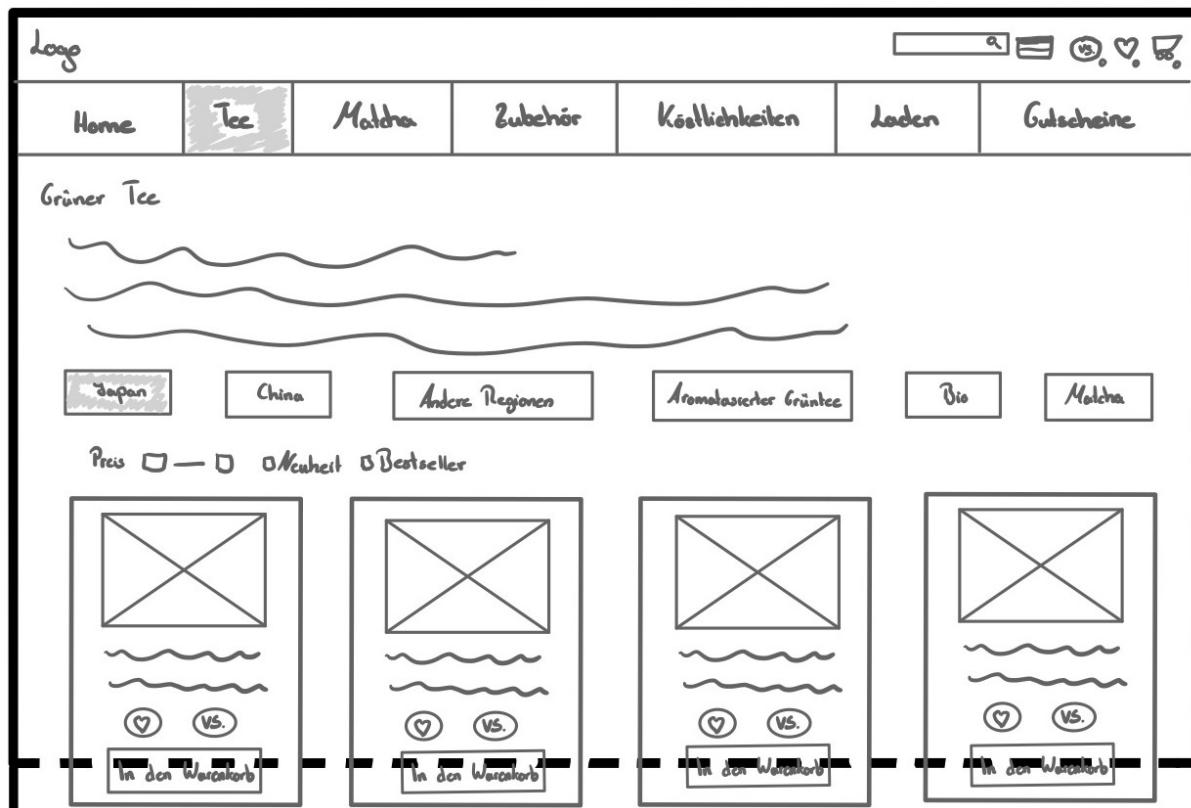


Abbildung 67. Entwurf des Online-Shops für Tee

Im Gegensatz zu der Seite „Köstlichkeiten“ sind in dem Online-Shop für Tee Filtermöglichkeiten integriert sowie die Möglichkeit die Produkte bereits in der Übersicht der Vergleichsliste, dem Merkzettel oder dem Warenkorb hinzuzufügen.

Die Gestaltungsideen für die Seite „Home“ und „Alles über Tee“, die Produktseiten für die Tees sowie das Menü wurden übernommen und weiter ausgearbeitet. Zudem entstanden erste Entwürfe für die Vergleichsseite und den Warenkorb. Die weiteren gezeichneten Wireframes befinden sich im Anhang und unterscheiden sich bezüglich Elementen wie den Header nicht.

Die Methode war besonders hilfreich, um erste Ideen zu visualisieren und diese dadurch besser bewerten zu können. Zudem ergab sich ein gutes Grundgerüst für den darauffolgenden Low-Fidelity-Prototypen.

### 3.1.3.6 Styleguide

Der Styleguide wurde in Figma entwickelt und legt Gestaltungsrichtlinien für die manuell gestaltete Anwendung fest. Er definiert die Farben der Buttons, Icons, Überschriften, Texte und den Hintergrund der Anwendung. Weiterhin wird die Schriftart, die Größe und die Stärke von Überschriften, Texten und Buttons festgelegt. Auch die Schriftart, der Zeilenabstand, die Größe, die Stärke und die Farbe der Schrift und der Icons im Logo werden geregelt. Zudem sind die Größen und die Farben der einzelnen Icons festgelegt. Zusätzlich sind auch die einzelnen Elemente wie der Hintergrund der Produkte oder die Sternenbewertung bezüglich der Größe und der Farbe genau vorgegeben. Das Aussehen des Headers und des Footers wird ebenfalls bezüglich der Farbe und der Größe genau definiert.

Der Styleguide war eine wertvolle Hilfe, um Design-Entscheidungen übersichtlich zu dokumentieren und sicherzustellen, dass das gesamte Erscheinungsbild der Anwendung einheitlich und stimmig bleibt.

## Farben



Abbildung 68. Farben der Hintergründe

Für den Hintergrund aller Seiten wurde ein neutrales Weiß mit einem leichten Beigeton gewählt. Dies verhindert eine visuelle Überladung und ermöglichte mehr Gestaltungsfreiheit bei den Buttons. Die Produktelemente heben sich mit einem dezenten Hellgrün vom Hintergrund ab, eine Farbe, die nicht zu grell ist und sich harmonisch ins Gesamtbild der Website einfügt. Um dem Nutzer ein besseres Feedback zu geben, wurde ein Hover-Effekt für die Produkte integriert. Beim Hovern über ein Produkt ändert sich dessen Farbe in ein kräftigeres Grün, das sich gut abhebt, aber dennoch eine gute Lesbarkeit der schwarzen Schrift gewährleistet.



Abbildung 69. Produkt ohne und mit Hover-Effekt

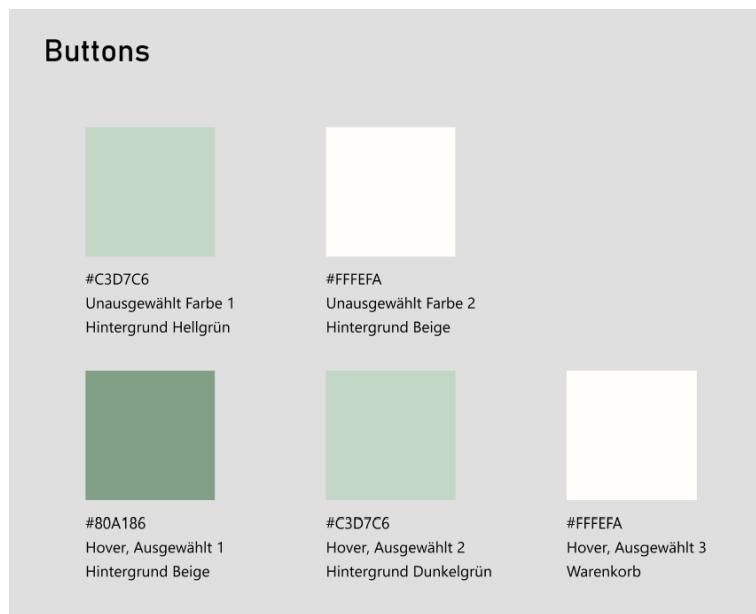


Abbildung 70. Farben der Buttons

Für die Buttons wurden aufgrund unterschiedlicher Hintergründe zwei Farbvarianten entwickelt. Buttons auf weißem Hintergrund sind in einem hellen Grün gehalten und wechseln beim Hovern oder Klicken zu einem dunkleren Grün für eine bessere Sichtbarkeit. Der Hover-Effekt wurde eingerichtet, um dem Nutzer ein besseres Benutzererlebnis durch ein Feedback der Buttons zu ermöglichen.

Mehr erfahren

Mehr erfahren

Abbildung 71. Button ohne und mit Hover-Effekt

Buttons auf einem hellgrünen Hintergrund, wie bei den Produktelementen, sind weiß mit einem leichten Beigeton und wechseln beim Hovern oder Klicken zu einem hellen Grün, um sich vom dunkleren Hover-Hintergrund der Produktelemente abzuheben.



Abbildung 72. Produkt ohne und mit Hover-Effekt bei dem Button

Die Buttons im Warenkorb benötigen eine spezielle Farbvariante, da der Hintergrund des Produktvorschlags zunächst weiß ist. Um sich gut abzuheben, ist der Button in diesem Zustand hellgrün. Beim Hovern wechselt der Hintergrund jedoch zu einem dunklen Grün, weshalb der Button dann eine weiße, leicht beige Farbe erhält, um weiterhin sichtbar zu bleiben.



Abbildung 73. Warenkorb Produkt ohne und mit Hover-Effekt

Auch die Buttons und andere interaktive Elemente wie die Produktelemente besitzen abgerundete Formen und Schlagschatten, um sie vom Hintergrund abzuheben und eine dezente Tiefenwirkung zu erzeugen.



Abbildung 74. Farben der Schriften

Die Überschriften der ersten und zweiten Ebene sind in dunkelgrün gehalten, um sich deutlich von den anderen Elementen abzuheben. Die dritte Überschrift die Button-Beschriftungen sowie die regulären Texte sind in schwarz gehalten, um eine optimale Lesbarkeit zu gewährleisten.



Abbildung 75. Farbe der Icons

Alle Icons sind einheitlich dunkelgrün, insbesondere im Header, wo sie sich von den anderen Elementen gut abheben. Die Icons sind inhaltsbezogen und so gestaltet, dass sie weitgehend ohne Schrift auskommen. Die meisten Icons haben eine einheitliche, dunkelgrüne Farbe und abgerundete Formen. Zudem wurde ein leichter Schlagschatten hinzugefügt, um sie vom Hintergrund abzuheben und eine dezentre Tiefenwirkung zu erzeugen.



Abbildung 76. Icons im Header



Abbildung 77. Icons unter Home

## Typographie

**Typographie**

<b>ABCdefg</b>	<b>ABCdefg</b>	<b>ABCdefg</b>
Überschrift 1 Schriftart Bahnschrift Größe 40 Stärke Regular Farbe #004511	Überschrift 2 Schriftart Bahnschrift Größe 32 Stärke Regular Farbe #004511	Überschrift 3 Schriftart Segoe UI Größe 20 Stärke Semibold Farbe #000000
<b>ABCdefg</b>	<b>ABCdefg</b>	<b>ABCdefg</b>
Text Schriftart Segoe UI Größe 16 Zeilenabstand 25 Stärke Regular Farbe #000000	Button Schriftart Segoe UI Größe 14 Stärke Regular Farbe #000000	

Abbildung 78. Typographie

Für die Website wurden zwei verschiedene seriflose Schriften gewählt, die einen modernen und cleanen Look vermitteln und besonders auf Bildschirmen gut lesbar sind. Um die erste und zweite Überschrift optisch hervorzuheben, wurde neben einer anderen Farbe auch eine abweichende Schriftart verwendet. Diese unterscheidet sich jedoch nur leicht, insbesondere in der Stärke, um das Gesamtbild

harmonisch und ruhig zu halten. Zudem variieren die Überschriften der ersten und zweiten Ebene in ihrer Schriftgröße, um eine klare Hierarchie zu schaffen.

Die dritte Überschrift sowie die Texte und Button-Beschriftungen verwenden eine einheitliche Schriftart, die sich jedoch in Größe und Stärke unterscheidet. Die dritte Überschrift hat eine Schriftgröße von zwanzig und die Stärke „Semibold“. Reguläre Texte sind mit einer Schriftgröße von sechzehn und Buttons mit einer Schriftgröße von vierzehn in der Stärke „Regular“ gehalten, um eine stimmige und gut lesbare Typografie zu gewährleisten.

## Logo

Das Logo wurde im gleichen Designstil wie die Icons gestaltet, um ein einheitliches und stimmiges Gesamtbild zu schaffen. Zur Entwicklung eines passenden Logos wurden zunächst verschiedene Ideen per Hand in Good Notes mit einem Apple Pencil skizziert. Dabei kamen typische Elemente rund um das Thema Tee, wie Teebeutel, Teetassen, Teekannen und Teeblätter, zum Einsatz. Zudem wurde ein Herz mehrfach integriert, inspiriert durch den Namen des Ladens „Tee Liebe Erlangen“. Die final ausgewählte Skizze wurde grün eingekreist, da sie die wichtigsten Elemente passend zum Laden und dessen Namen vereinte.

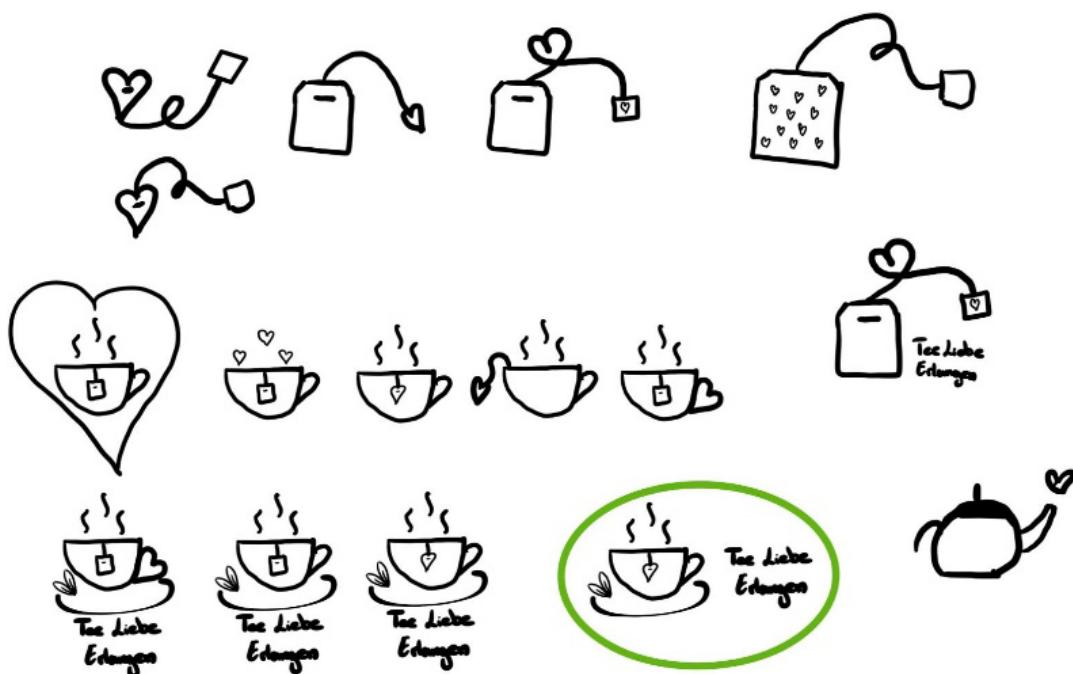


Abbildung 79. Erste Skizzen des Logos

Anschließend wurde das Logo für den ersten High-Fidelity-Prototyp ausgearbeitet und farblich in hellen sowie dunkelgrünen Tönen gestaltet, um es optimal an das Design der Website anzupassen. Die Schriftart war hier zunächst anders als die restlichen Schriftarten. Ursprünglich wurde eine geschwungene, handschriftlich anmutende Schriftart gewählt, um dem Logo eine persönliche Note zu verleihen.



Abbildung 80. Logo des High Fidelity Prototyp 1

Für den High-Fidelity-Prototyp wurde das Logo jedoch überarbeitet. Die Teeblätter und die Untertasse wurden entfernt, um das Logo klarer und großflächiger darstellen zu können. Zudem wurde die Schriftart an die Typografie der ersten und zweiten Überschriften angepasst, um ein harmonisches Gesamtbild zu gewährleisten. Diese Änderungen ergaben sich auch aus den Nutzerfeedbacks der durchgeföhrten Tests, die eine klarere und einheitlichere Gestaltung forderten.

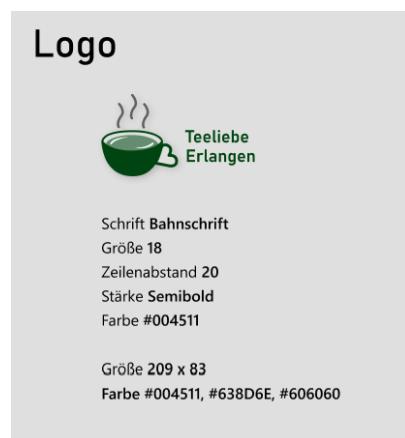


Abbildung 81. Finales Logo

## Header



Abbildung 82. Header

Im Header wurden zahlreiche zuvor getroffene Design-Entscheidungen umgesetzt, um ein stimmiges und einheitliches Erscheinungsbild zu schaffen. So dominieren bereits verwendete Farben wie helles und dunkles Grün, während die einzelnen Menüpunkte durch sanfte Rundungen und einen dezenten Schlagschatten hervorgehoben werden. Besonderes viel Wert lag auf der Übersichtlichkeit, da diese ein zentraler Kritikpunkt in der Bewertung der Mitbewerber war. Um dies zu gewährleisten, wurden ausreichende Abstände zwischen den Elementen eingehalten und ausschließlich die wichtigsten Inhalte im Header platziert, um eine visuelle Überladung zu vermeiden.

## Footer

Kontakt:	Sie finden uns auch bei:	Shop Service:	Informationen:	Newsletter:
4564 654646546 TeaLove.Erlangen@gmx.de <a href="#">Kontaktformular</a>	<a href="#">facebook.com/TeaLove</a> <a href="#">instagram.com/TeaLove</a> <a href="#">pinterest.com/TeaLove</a> <a href="#">https://youtube.com/TeaLove</a>	<a href="#">Cookie-Einstellungen</a> <a href="#">Zahlungsarten</a> <a href="#">Versand, Lieferung &amp; Retoure</a> <a href="#">Impressum</a> <a href="#">Datenschutz</a> <a href="#">AGB</a> <a href="#">FAQ-Häufige Fragen</a>	<a href="#">Standort &amp; Öffnungszeiten</a> <a href="#">Kontakt</a> <a href="#">Über den Laden</a> <a href="#">Über Uns</a> <a href="#">Aktuelle Stellenangebote</a>	Melden Sie sich jetzt zu unserem Newsletter an und erfahren Sie alle Neuigkeiten rund um unseren Laden, Angebote und hilfreiche Informationen rund um Tee! Wir schenken Ihnen zudem einen 10% Code auf Ihre erste Bestellung!  <a href="#">Jetzt anmelden!</a>
Obere Karlstraße 30 91054 Erlangen				
Komm vorbei! Montag - Donnerstag 07:00-12:00 & 13:00-19:00 Uhr Samstag 10:00-13:00 & 14:00-20:00 Uhr				

Abbildung 83. Footer

Auch der Footer folgt den Designentscheidungen der restlichen Website und wurde bewusst minimalistisch gestaltet, um eine optimale Übersichtlichkeit zu gewährleisten.

### 3.1.3.7 Low-Fidelity-Prototype

Der Prototyp wurde in Figma basierend auf den Methoden Brainstorming, Future User Journey Map und den Wireframes erstellt.

Dabei lag der Fokus vorrangig auf der Funktionalität sowie dem Aufbau der Anwendung und weniger auf den Designentscheidungen oder den konkreten Inhalten. Deshalb wurden überwiegend Platzhalter und Fülltexte verwendet. Jedoch wurden Überschriften, Buttons und Filter sowie der Header und der Footer bereits mit den korrekten Inhalten gefüllt, um die Nutzung für die späteren Usability-Tests verständlicher zu gestalten.

Dabei gilt es auch zu erwähnen, dass von einer weiteren Unterteilung des Matchapulvers nach Anbaugegenden oder Geschmacksrichtungen, wie bio, mild, würzig und süß, bewusst verzichtet wurde. Eine solche Unterteilung war weder in der fiktiven Auftragsstellung enthalten noch stand genügend Bildmaterial im Internet zur freien Verfügung. Zudem war eine weitere Unterteilung nicht relevant für das Abtesten der Forschungsfragen, da der Fokus auf die Benutzerfreundlichkeit und -erfahrung liegt und nicht auf dem Produktangebot.

In den anschließenden User Tests, die mit vier Testpersonen durchgeführt wurden, von denen zwei UI/UX-Experten waren und zwei Laien in dem Bereich, ergaben sich viele neue Erkenntnisse, die dazu führten, den Prototyp weiter zu verbessern. Die Erkenntnisse aus den Tests wurden in Miro mithilfe des Feedback Capture Grids auf Basis der Beobachtungen während der Tests sowie aus den anonymen Umfragen festgehalten. Dabei wurden sowohl positive Aspekte, die beibehalten oder weiter ausgebaut werden sollten, als auch negative Aspekte, die verbessert oder ersetzt werden mussten, festgehalten.

Positive Punkte, die gut funktionierten, waren zum einen der Jump Up Button, die Vergleichsliste, das Berechnen der Versandkosten, der Umgang mit dem Menü und der übersichtliche Footer. Zum anderen waren die Sortierung, die Filtermöglichkeiten vor allem auch nach Preisen und der Merkzettel gut und sinnvoll. Die Ladeninformationen konnten schnell gefunden werden, unter anderem auch durch die kurzen aber prägnanten Informationen auf Home und der Möglichkeit ausführlichere Informationen durch einen Klick auf den Button „Mehr erfahren“ zu erhalten. Zudem wurden die Verlinkungen unter Standort und Öffnungszeiten positiv bewertet. Weiterhin wurden die Informationen und die Auswahlmöglichkeiten

in den Produktseiten als hilfreich und ansprechend wahrgenommen. Auch die Anzeige des Menüs, wo man sich befindet, wurde positiv bewertet sowie die Platzierung des Warenkorbes und des Buttons zum Umstellen der Sprache über dem Menü. Zuletzt wurde auch der Warenkorb und der Weg dahin positiv bewertet, da zum einen der Warenkorb Pop-up verständlich und gut strukturiert ist und zum anderen Funktionen wie das Berechnen der Versandkosten hilfreich sind. Auch die Info, ab wann die Bestellung versandkostenfrei ist, und der Vorschlag von Produkten, die dabei helfen dieses Ziel zu erreichen, wurden als positiv empfunden.

What worked					
Merkzettel gut und sinnvoll	Verlinkungen unter Standort und Öffnungszeiten sind gut	schnelles Auffinden der Ladeninformationen	<b>Jump Button</b>	Gute Sortierung	Filtermöglichkeiten sind hilfreich
Knappe Infos auf Home -> Entscheidung für mehr Infos möglich	Gut, dass man nach Preisklasse filtern kann	<b>Umgang mit dem Menü</b>	Infos und Auswahlmöglichkeiten in den Produktseiten gut	Warenkorb und Weg dahin verständlich	Vergleichsliste
Info ab wann es Versandkostenfrei ist + Produkte	Berechnen der Versandkosten	warenkorb PopUp verständlich und gut strukturiert	Button im Menü oben rechts gut aufgefunden	Menü zeigt an wo man sich befindet	übersichtlicher Footer

Abbildung 84. Feedback Capture Grid für den Low-Fidelity-Prototype – What worked

Negative Punkte, die verbessert werden müssen, waren zum einen, dass die Ladeninformationen in Home nicht einzeln anklickbar sind und man auf die entsprechende Unterseite weitergeleitet wird und zum anderen, dass diese noch kompakter werden müssen. Weiterhin war den Testpersonen vereinzelt nicht klar, was in der Vergleichsliste verglichen wird. Im Warenkorb wurde kritisiert, dass die Anzeige, ab wann es versandkostenfrei ist und die Vorschläge für weitere zu kaufende Produkte, um dieses Ziel zu erreichen, zu groß ist und nerven würden. Beim Zubehör wurde auf einen Fehler in der Preisangabe hingewiesen, da hier die Angabe Preis pro Kilogramm keinen Sinn ergab. Zudem sollte bei den Filtern für einen aromatisierten Tee die Benennung bei allen Teesorten gleich sein, da bisher bei dem Grüntee die Benennung „aromatisierter Grüntee“ und bei dem Oolong Tee „aromatisierte Tees“ lautet und dies nicht einheitlich ist und den Nutzer verwirrt. Auch die Benennung des Menüpunktes verunsicherte die Nutzer, da sie hier nicht wussten, was genau sie erwartete und sie zunächst nicht verstanden, dass dies nur ein Angebot vor Ort ist. Weiterhin wurde das Fehlen einer Rücktaste kritisiert, jedoch wäre diese automatisch in einem Browser verfügbar und muss in der Anwendung selbst nicht integriert werden.

Jedoch könnte es dem Nutzer vereinfacht werden einen Schritt in der Anwendung zurückzugehen. Zuletzt beschwerte sich eine Testperson über das lange Laden sowie über ein Aufhängen der Anwendung. Dies ist jedoch dem Programm oder der Internetverbindung der Testperson geschuldet. Somit ist eine Verbesserung bei dem High-Fidelity-Prototypen in diesem Bereich nicht möglich.



Abbildung 85. Feedback Capture Grid für den Low-Fidelity-Prototype – What could be improved

Die Testpersonen gaben jedoch auch Ideen für eine Verbesserung der Anwendung an. Zum einen könnte der Menüpunkt „Köstlichkeiten“ in „Köstlichkeiten vor Ort“ umbenannt werden, damit der Inhalt für den Nutzer besser verständlich ist. Zum anderen sollte bei der Vergleichsmöglichkeit die Vergleiche auf ihre Sinnhaftigkeit erklärt werden, damit dies auch Anfänger in dem Bereich Tee besser verstehen.

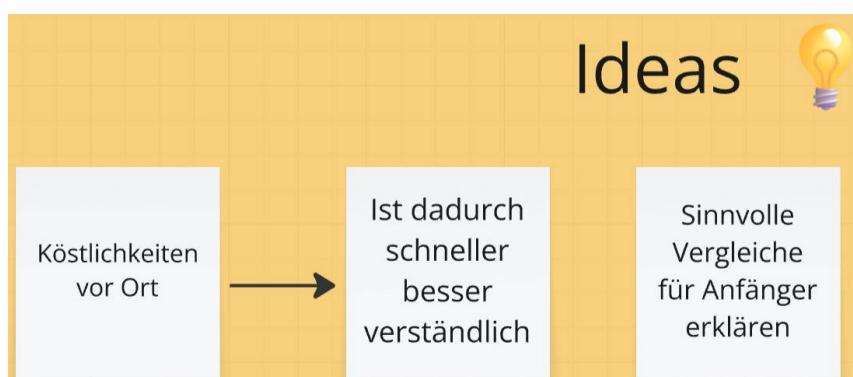


Abbildung 86. Feedback Capture Grid für den Low-Fidelity-Prototype – Ideas

Die Erstellung des Low-Fidelity-Prototypen, die Usability-Tests sowie die anschließende Dokumentation der Ergebnisse in dem Feedback Capture Grid half sehr dabei die weiteren Anforderungen und Wünsche der Nutzer an die neue Anwendung herauszufinden und festzuhalten. Diese Erkenntnisse bilden nun die Grundlage für die Umsetzung des High-Fidelity-Prototyps.

### 3.1.3.8 High-Fidelity-Prototype 1

Der Prototyp wurde in Figma erstellt, basierend auf den Methoden Brainstorming, Wireframes und dem Low-Fidelity-Prototypen. Zusätzlich wurde die Informationsarchitektur genutzt, um einen Überblick über die notwendigen Inhalte und Verlinkungen zu behalten.

Der High-Fidelity-Prototyp stellt die erste ausgearbeitete Version der Anwendung dar, insbesondere in Bezug auf die Designentscheidungen und den Inhalten. Weiterhin wurden weitere Seiten angelegt und die notwendigen Verlinkungen hergestellt. Auch das Feedback aus den Usability-Tests des Low-Fidelity-Prototypen wurde umgesetzt. Anschließend entstand eine umfassende Anwendung, mit der erneut Usability-Tests durchgeführt und ein Feedback eingeholt werden konnten, um den Prototypen weiter zu verbessern. Hier wurden die Tests mit fünf Testpersonen durchgeführt von denen zwei UI/UX-Experten waren und drei Laien in dem Bereich. Die Erkenntnisse wurden erneut in einem Feedback Capture Grid in Miro dokumentiert, basierend auf den Beobachtungen während der Tests sowie auf den anonymen Umfragen. Auch hier gab es sowohl positive Punkte, die beibehalten oder weiterentwickelt werden sollten, als auch negative Aspekte, die verbessert oder überarbeitet werden mussten.

Positive Punkte, die gut funktionierten, waren zum einen, dass die Schrift gut lesbar war, die Bilder und Texte ein gutes Bild des Ladens wiedergaben, die Filterauswahl, die Größe und das Design der Buttons, der Hovereffekt und die passenden Farben. Zum anderen konnten die Testpersonen gut mit dem Warenkorb und dem Pop-up umgehen sowie mit den Inhalten und Funktionen der Vergleichsliste, die auch gut auffindbar war. Auch die Inhalte in Home waren passend. Weiterhin wurde die Strukturierung des Menüpunktes „Genuss vor Ort“ als positiv bewertet. Die Menüstruktur allgemein wurde lobend erwähnt. Vor allem wurde das Auffinden der Ladeninformationen, die Teeauswahl und die Teedosenauswahl als positiv bewertet. Zuletzt wurde auch das Zurückleiten des Logos zu Home gelobt.



## What worked

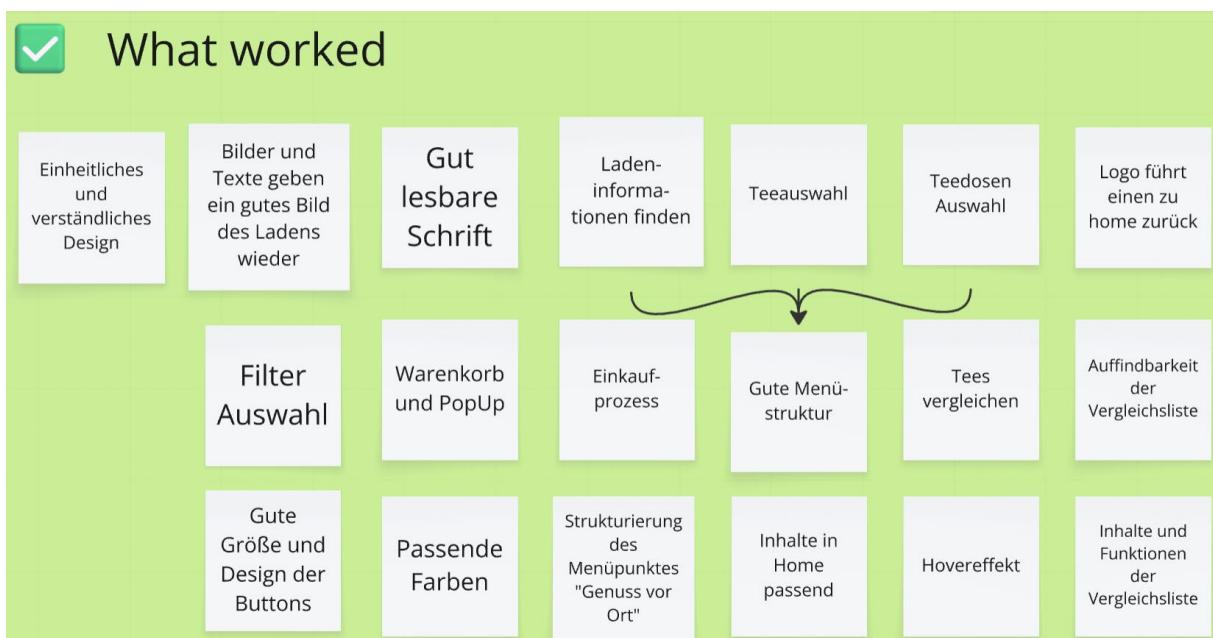


Abbildung 87. Feedback Capture Grid für den High-Fidelity-Prototype – What worked

Die negativen Punkte, die verbessert werden müssen, wurden jedoch in die Kategorien wird nicht umgesetzt, wird umgesetzt und wird eventuell umgesetzt sortiert. Ein Kritikpunkt, der auf jeden Fall verbessert werden muss, war zum einen die Position des „Kategorie hinzufügen“ Buttons in der Vergleichsliste, der zu weit von den Kategorien entfernt war und dessen Funktion dadurch nicht erkannt wurde. Zum anderen sollte die Motivanzeige bei den Gutscheinen verbessert werden, da diese bisher zu schlecht lesbar ist. Auch das Öffnen des Untermenüs durch die Pfeile muss verbessert werden, da die Testpersonen nicht verstanden, dass auf den Pfeil geklickt werden muss, um das Menü weiter zu öffnen. Weiterhin soll auf der Seite „Alles über Tee“ eine Möglichkeit eingerichtet werden, wieder zurück zu den anderen Tee Seiten gehen zu können. Zudem sollten die Untermenüs eine Rundung bekommen. Einige Testpersonen haben die Funktion, die Versandkosten im Warenkorb zu berechnen, mit der Dateneingabe verwechselt. Daher muss dies erkenntlicher gemacht werden. Das Logo und vor allem die Schrift sollte verbessert werden. Zuletzt sollte es ermöglicht werden die Bilder vor allem unter Ladeninformationen größer anzuschauen.

Ein Kritikpunkt, der jedoch nur eventuell umgesetzt wird, war zum einen, dass beim Hovern über Buttons eine Beschreibung erscheinen soll, um vor allem bei den Buttons im Header, wie bei der Vergleichsliste, die Funktion deutlicher zu machen. Zum anderen wurde kritisiert, dass bei den Gutscheinen zu viele Buttons angezeigt werden, um den Wert auszuwählen. Weiterhin wurde das Design als flach und nicht modern bewertet und auch die Home Seite wurde als zu minimalistisch angesehen. Auch die Kopfzeile wurde kritisiert, da diese noch nicht stark genug hervortritt. Zuletzt sollte zum Beispiel auf der Seite „Teegeschirr“ oder den Unterseiten der Teesorten eine Möglichkeit eingerichtet werden wieder zu der Übersichtsseite zurückzukehren. Diese Vorschläge müssen erst auf Sinnhaftigkeit und mögliche Umsetzung in der Anwendung getestet werden. Es gab keine Kritikpunkte deren Umsetzung von vornherein komplett ausgeschlossen wurde.

What could be improved					
Position des "Kategorie Hinzufügen"-Button	Beschreibung der Buttons beim drüber Hovern	Motivanzeige bei den Gutscheinen besser lesbar machen	Öffnen des Untermenüs durch die Pfeile verbessern	Zu viele Buttons bei den Gutscheinen	Design wirkt flach und nicht modern
Im Teegeschirr / Grünen Tee wieder zurückgehen	In Alles über Tee wieder zurück gehen	Rundung für die Untermenüs	Home sehr minimalistisch	Versandkosten berechnen mit Datenangabe verwechselbar	Kopfzeile könnte noch stärker hervortreten
Logo minimalisieren	Font verbessern	Bilder größer anschauen			

Abbildung 88. Feedback Capture Grid für den High-Fidelity-Prototype – What could be improved

Die Testpersonen gaben jedoch auch Ideen für eine Verbesserung der Anwendung an. Diese wurden jedoch auch in die Kategorien wird nicht umgesetzt, wird umgesetzt und wird eventuell umgesetzt sortiert. Ideen, die umgesetzt werden, waren zum einen, dass das Untermenü durch Hovern über den Pfeil mit einer kleinen Zeitverzögerung geöffnet wird. Zum anderen soll der „Kategorie hinzufügen“ Button in der Vergleichsliste näher an den Kategorien zum Vergleichen der Tees positioniert werden. Weiterhin soll bei Elementen wie den Buttons ein Schlagschatten hinzugefügt werden. Zuletzt soll es Empfehlungen für unerfahrene Teetrinker geben, um die Teefindung zu erleichtern.

Ideen, die jedoch nur eventuell umgesetzt werden, waren zum einen, dass die Sprünge des auszuwählenden Wertes der Gutscheine größer werden sollen. Zum anderen soll es unter Home einen Hinweis auf die Vergleichsliste in Form eines Banners geben. Auf Home soll es zudem mehr Bilder geben. Die letzte Idee bestand aus einem Quiz, mit dem der Kunde herausfinden kann, welcher Tee für die individuellen Ansprüche der Beste ist. Diese Vorschläge müssen erst wieder auf Sinnhaftigkeit und mögliche Umsetzung in der Anwendung getestet werden.

Es gab dabei keine Ideen deren Umsetzung von vornherein komplett ausgeschlossen wurde.



Abbildung 89. Feedback Capture Grid für den High-Fidelity-Prototypen – Ideas

Die Erstellung des ersten High-Fidelity-Prototypen, die darauffolgenden Usability-Tests sowie die Erfassung der Ergebnisse in dem Feedback Capture Grid halfen sehr dabei, die weiteren Anforderungen und Wünsche der Nutzer an die neue Anwendung herauszufinden und festzuhalten. Diese Erkenntnisse fließen nun in die Entwicklung eines zweiten High-Fidelity-Prototyps ein.

### 3.1.3.9 High-Fidelity-Prototype 2

Der zweite High-Fidelity-Prototyp wurde ebenfalls in Figma entwickelt, basierend auf dem ersten High-Fidelity-Prototypen und den Erkenntnissen aus dem Feedback Capture Grid. Zusätzlich wurden weitere Methoden wie die User Stories, die „How Might we...?“ Fragen, das Brainstorming und die Future-State Journey Map eingesetzt. Diese dienten als Erinnerung an das Ziel der Arbeit und stellten sicher, dass die Wünsche und Anforderungen der Nutzergruppe auch in diesem finalen Schritt berücksichtigt wurden. Um erneut einen Überblick über die notwendigen Inhalte und Verlinkungen zu erhalten, wurde zudem die Informationsarchitektur verwendet.

Der zweite High-Fidelity-Prototyp stellt eine verbesserte Version des ersten High-Fidelity-Prototypen dar, in dem insbesondere das Feedback aus den Usability-Tests des ersten Prototyps umgesetzt wird. Anschließend entstand eine verbesserte Anwendung, die nun in weiteren Usability-Tests evaluiert und mit einer KI-generierten Anwendung verglichen werden kann.

Die Usability-Tests wurden mit insgesamt zweiundvierzig Testpersonen durchgeführt. Jeweils eine Testperson testete eine der beiden Anwendungen. Das heißt einundzwanzig testeten die KI-generierte Anwendung und einundzwanzig testeten die manuell erstellte Anwendung. Um Verzerrungen oder Beeinflussungen zu vermeiden, wussten die Testpersonen während des Tests nicht, welche Anwendung sie testen. Die Testergebnisse basieren auf den Beobachtungen und den Notizen während der Tests sowie aus einer anonymen Umfrage zu der Anwendung. Der genaue Ablauf der Tests und dessen Ergebnisse werden in Kapitel [3.3 Testen der Lösungen gegen die Nutzungsanforderungen](#) und Kapitel [4. Ergebnisse des Usability-Tests](#) genauer behandelt.

Bei der Erstellung beider High-Fidelity-Prototypen wurden zudem Bilder von pixabay<sup>14</sup> für die Produkte und einen Ausschnitt von Google Maps<sup>15</sup> für den Standort verwendet.

## **3.2 Erstellung der KI-generierten Applikation**

Im Folgenden werden die notwendigen Schritte zur Vorbereitung der Erstellung vorgestellt sowie die anschließende Erstellung der Applikation durch einen neutralen Laien.

### **3.2.1 Programmauswahl**

Zunächst wurde im Internet nach Anbietern recherchiert, die einen KI-gestützten Website-Builder anbieten, um eine Übersicht zu bekommen und die geeignete Auswahl treffen zu können. Nach der Recherche ergab sich eine kleine Auswahl an Anbietern, darunter WIX, IONOS, Durable, uizard und Hostinger.

Die Auswahlkriterien umfassten zum einen natürlich einen KI-gestützten Website-Builder und zum anderen ein gutes Preis-Leistungs-Verhältnis, eine benutzerfreundliche Bedienung, umfassende Funktionen, die Möglichkeit die Eingaben und die Website auf Deutsch zu erstellen und viele informative Tutorials im Internet, zur Einarbeitung des Erstellers in die Funktionen des KI-gestützten Website-Builders.

Nach Anwendung dieser Kriterien blieben die Anbieter WIX, Durable und IONOS in der engeren Auswahl. Die anderen Anbieter wurden verworfen, da essenzielle Funktionen entweder nur gegen zusätzliche Kosten verfügbar waren, die bei den verbleibenden Anbietern bereits im Paket enthalten waren. Weiterhin war nicht ersichtlich, ob die Sprache während der Erstellung und auf der finalen Website umgestellt werden kann. Dies hätte die Vergleichbarkeit mit der manuell gestalteten Website maßgeblich behindert.

Schließlich wurde der KI-gestützte Website-Builder von WIX ausgewählt, um die KI-Anwendung zu generieren. Die ausschlaggebendste Grund war die kostenlose Nutzung des Website-Builders, da nur optionale Erweiterungen, die für diese Arbeit irrelevant sind, wie eine individuelle Domain, kostenpflichtig sind. Dadurch hatte dieser Anbieter das beste Preis-Leistungs-Verhältnis. Weiterhin wurde der KI-gestützte Website-Builder im Vorfeld getestet und als sehr benutzerfreundlich empfunden, da die KI den Ersteller Schritt für Schritt durch den Prozess führt und klare Anweisungen gibt, die auch für Laien verständlich sind. Auch der Website Editor und die Einrichtung der Produkte ist übersichtlich und verständlich gestaltet. Besonders hilfreich ist hier zudem die Möglichkeit zwischen einer Chat gesteuerten Erstellung und einer Schritt für Schritt Eingabe aller wichtigen Informationen und Inhalte wählen und somit

---

<sup>14</sup> <https://pixabay.com/>

<sup>15</sup> [https://www.google.com/maps/place/Ob.+Karlstra%C3%9Fe,+91054+Erlangen/@49.5965495,11.0061401,571m/data=!3m2!1e3!4b1!4m6!3m5!1s0x47a1f8dfda447fe7:0xcd177eca70369d21!8m2!3d49.5965461!4d11.008715!16s%2Fg%2F1tz94rk0?entry=ttu&g\\_ep=EgoYMDI1MDQzMCA4wIKXMDSoASAFQAw%3D%3D](https://www.google.com/maps/place/Ob.+Karlstra%C3%9Fe,+91054+Erlangen/@49.5965495,11.0061401,571m/data=!3m2!1e3!4b1!4m6!3m5!1s0x47a1f8dfda447fe7:0xcd177eca70369d21!8m2!3d49.5965461!4d11.008715!16s%2Fg%2F1tz94rk0?entry=ttu&g_ep=EgoYMDI1MDQzMCA4wIKXMDSoASAFQAw%3D%3D)

selbst entscheiden zu können, welche Form der Erstellung am besten ist. Zudem gibt es im Internet viele umfangreiche, informative und verständliche Tutorials, die den KI-gestützten Website-Builder sehr gut erklären und die Testperson gut auf die Erstellung vorbereiten.

Die anderen Anbieter IONOS und Durabel wurden verworfen, da IONOS kostenpflichtig ist, jedoch keinen erkennbaren Mehrwert bezüglich der vorhandenen Funktionen für diese Arbeit bietet. Der Anbieter Durable wurde nicht ausgewählt, da dieser wenig Eingabemöglichkeiten zu Beginn der Erstellung bietet. Im Erstellungsprozess von WIX können mehr Eingaben getätigt werden, die für die Website wichtig sind und nicht fehlen dürfen. Somit wurde der Anbieter WIX für die Erstellung der KI-generierten Anwendung ausgewählt.

### **3.2.2 Erstellung**

#### **3.2.2.1 Kriterien für die Erstellung**

Die KI-erstellte Anwendung wurde unter der Anleitung einer neutralen Person ohne Erfahrung im Webdesign erstellt. Die Person hatte kein persönliches Interesse am Ergebnis der Arbeit, um eine bewusste oder unbewusste Beeinflussung der Qualität der Anwendung weder positiv noch negativ zu vermeiden.

#### **3.2.2.2 Vorbereitung der Erstellung**

Wichtig war, dass dem Ersteller eine thematische Übersicht und Regeln vorgegeben wurden, an der sich die Testperson richten und halten musste. Es musste vor allem festgelegt werden, was der Ersteller nach der ersten Version der KI verändern darf.

Die erste Regel erlaubte es nur das zu ändern, was mit KI verändert werden kann. Das heißt Bilder dürfen durch neu generierte KI-Bilder ersetzt werden, ein Logo darf mit KI entwickelt werden und Unterseiten dürfen mit der KI neu generiert und neu angeordnet werden. Zuletzt dürfen Texte nur mit KI geschrieben werden.

Die zweite Regel besagt, dass fehlende Produkte nach der Vorlage der KI ergänzt werden dürfen, um eine Vergleichbarkeit beider Websites mit identischem Produktangebot zu gewährleisten.

Weiterhin musste festgelegt werden, was nicht verändert werden darf. Zum einen darf nicht die Struktur beziehungsweise die Informationsarchitektur der Website verändert werden. Das heißt der Ersteller darf keine Abschnitte oder Inhalte löschen oder verschieben. Zum anderen darf er nicht den Sinn der Abschnitte verändern. Dies gilt insbesondere für die Überschriften, deren Wortlaut angepasst werden darf, aber nicht deren inhaltliche Bedeutung.

Die Einhaltung dieser Regeln war entscheidend, da die Bewertung der Website nicht das Können der Testperson, sondern die Leistungsfähigkeit der KI widerspiegeln sollte. Aus diesem Grund war es wichtig, dass alle inhaltlichen, gestalterischen und strukturellen Entscheidungen von der KI getroffen wurden.

### **3.2.2.3 Ablauf der Erstellung**

Zunächst wird die Testperson begrüßt und in das Thema eingeführt. Dabei werden das Projekt sowie dessen Ziel kurz erläutert. Anschließend wird der Testablauf besprochen, sodass die Person weiß, was auf sie zukommt und an welcher Stelle im Prozess sie sich später befindet.

Daraufhin werden die Richtlinien vorgestellt, anhand derer die Applikation erstellt werden soll. Diese liegen der Testperson ausgedruckt vor, damit sie die Vorgaben eigenständig mitlesen und während der Erstellung jederzeit nachlesen kann.

Im nächsten Schritt wird die Art der Website vorgestellt. Die Testperson erhält detaillierte Informationen über die Struktur der Website sowie über die notwendigen Inhalte, Informationen und Produkte, die enthalten sein müssen. Auch diese Angaben werden der Testperson schriftlich zur Verfügung gestellt, um das Nachlesen zu erleichtern.

Anschließend beginnt die Erstellung der Website. Ein wesentlicher Bestandteil dieses Prozesses stellt die Einarbeitung in das Programm dar. Hierfür kann die Testperson aus einer kleinen Auswahl vorgeschlagener Tutorials wählen oder sich eigenständig geeignete YouTube-Videos suchen. Sie entscheidet selbst, wann sie sich ausreichend informiert fühlt.

Nach der Einarbeitung wird die Website von WIX im Internet geöffnet. Dabei ist zu beachten, dass die Testperson sich mit dem Account des Erstellers dieser Arbeit eingeloggt hat, um den vollständigen Zugriff auf die Website für den späteren Verlauf zu gewährleisten. Nun beginnt die eigentliche Entwicklung der Website. Die Testperson wählt zunächst einen Weg der Erstellung aus. Dabei muss sie sich zwischen einer Chat gesteuerten Erstellung und einer Schritt für Schritt Eingabe aller wichtigen Informationen und Inhalte entscheiden, bei der die Testperson die KI mit den notwendigen Informationen versorgen muss. Im weiteren Verlauf können diese Angaben, wenn notwendig, nochmal verfeinert werden. Anschließend kann die Website noch bezüglich der Produkte und weiteren Details angepasst werden.

Die Erstellung der Website mit dem KI-Builder von Wix gilt als abgeschlossen, wenn die Testperson zufrieden ist und keine weiteren Änderungen mit der KI möglich sind.

### **3.2.2.4 Beobachtungen während der Erstellung**

Während der Erstellung wurde die Testperson die meiste Zeit über beobachtet, um wertvolle Einblicke zu erhalten. Zunächst sah er sich das Tutorial „How To Build A Website in 16 Minutes (Wix AI Builder Tutorial 2024)“ an (Santrel Media, 2024). Danach fühlte er sich ausreichend informiert und empfand den KI-gestützten Website-Builder als verständlich aufgebaut. Auch das im Video gezeigte Ergebnis wirkte vielversprechend. Anschließend ließ er sich die Richtlinien sowie die erforderlichen Inhalte für die Website noch einmal durch. Nachdem er sich erneut einen Überblick verschaffen hatte, fühlte er sich bereit, mit der Erstellung der Website zu beginnen.

Zu Beginn musste der Ersteller sich für eine Form der Website-Erstellung entscheiden. Dabei standen zwei Optionen zur Verfügung. Eine Erstellung mit oder ohne Chat-Funktion. Er folgte der Empfehlung des Tutorials und entschied sich für die Erstellung ohne Chat-Funktion. Einige Schritte fielen der Testperson während der Erstellung leicht, insbesondere die Eingabe der Standardinformationen, die der KI-Builder

benötigte. Dazu gehörten Angaben zu der Art der zu erstellenden Website, der Name des Unternehmens, die bisherigen Erfahrungen mit einem Onlineshop sowie Angaben zu der Art der Produkte und wo diese angeboten werden sollen. Nach der Eingabe dieser Daten gelangte er zu einer Übersicht, die er schrittweise abarbeiten musste.

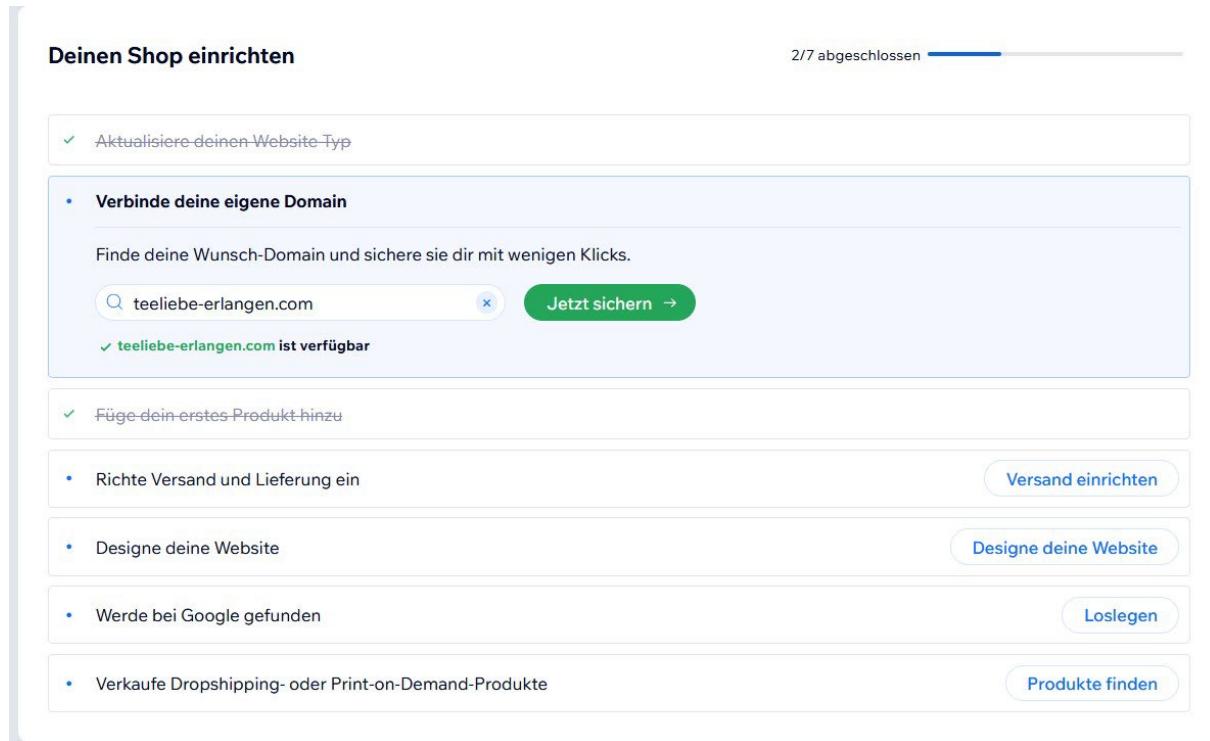


Abbildung 90. Abzuarbeitende Übersicht bei der Erstellung der KI-Applikation

Hierbei erinnerte er sich an das angesehene Tutorial und wusste, welche Angaben hier erforderlich waren. Als Erstes wollte er eine Wunschdomain angeben, stellte jedoch fest, dass dies mit zusätzlichen Kosten verbunden ist. Da die bisherige Domain für die Usability-Tests ausreichend sein sollte und später noch geändert werden kann, entschied er sich dazu diesen Schritt vorerst zu überspringen. Anschließend begann er mit der Erstellung der Produkte für die spätere Website. Nachdem er sich die vorgeschlagenen Einstellungen angesehen hatte, stellte er jedoch fest, dass die Eingabe aller Produktinformationen sehr zeitaufwendig wäre. Dies frustrierte ihn, da er am liebsten mit der eigentlichen Erstellung der Website fortfahren wollte. Daher kehrte er zu der Übersicht zurück und übersprang zunächst diesen Schritt.

Als Nächstes war die Einstellung der Versandoptionen an der Reihe. In diesem Schritt wurde die Testperson zunehmend frustrierter. Zwar unterstützt die KI in einem Chat bei der Auswahl der Länder und gab Tipps, legte die Länder jedoch nicht selbstständig an. Daher entschied er sich vorerst dazu, bei den bisherigen Angaben zu bleiben. Das heißt für einen kostenlosen Versand im In- und Ausland. Er behielt sich jedoch vor, diese Einstellungen eventuell später noch anzupassen.

Anschließend klickte er auf „Website designen“ und erhielt die Information, dass die Website nun von der KI erstellt wird. Bereits an dieser Stelle fiel ihm auf, dass der Erstellprozess und das Laden der Website länger dauert als erwartet. Trotzdem blieb er gespannt auf das Ergebnis. Als die Website

schließlich angezeigt wurde, war er zunächst wenig überzeugt, hoffte aber, dass sich im weiteren Prozess noch Verbesserungen ergeben werden.

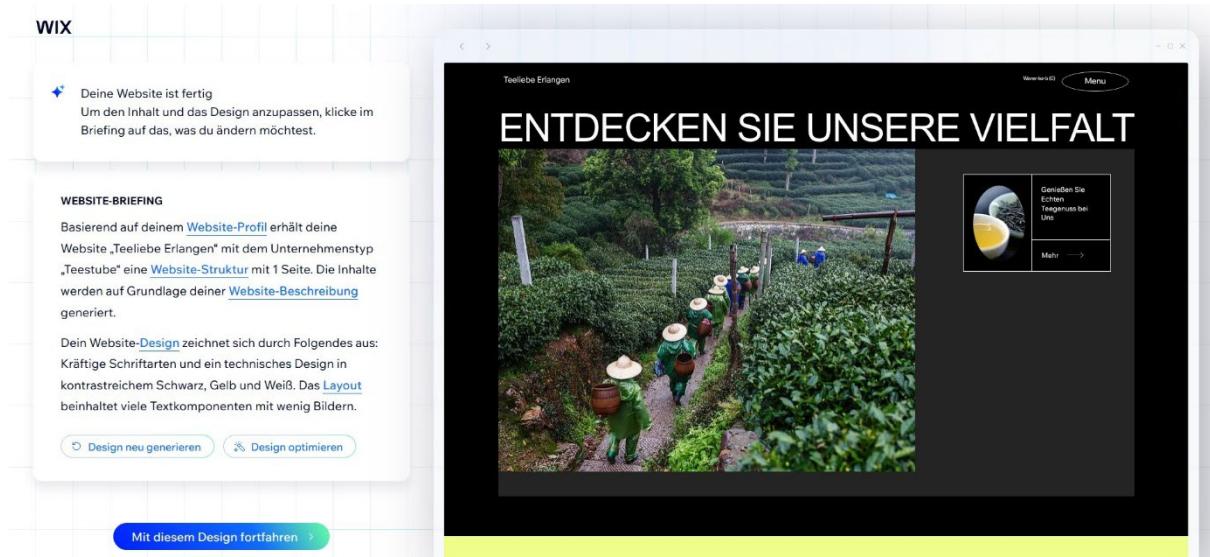


Abbildung 91. Erstes Design der KI-Applikation

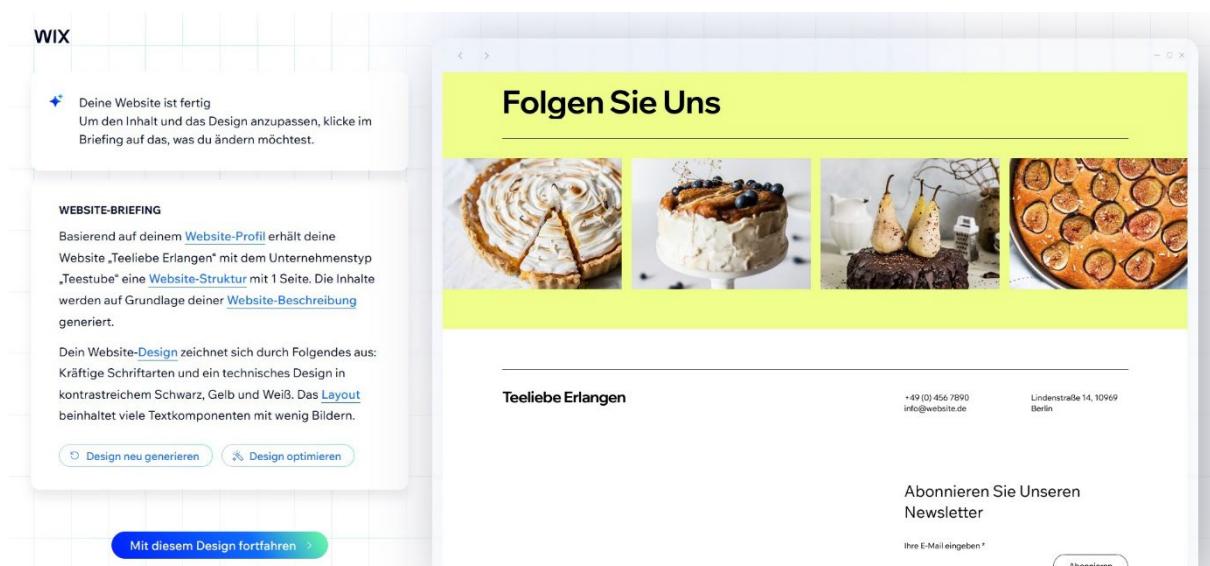


Abbildung 92. Weitere Ansicht des ersten Designs der KI-Applikation

Auf der linken Seite sah er einen Text und erinnerte sich aus dem Video daran, dass er die Website unter den blauen Links weiter anpassen kann. Zunächst öffnete er den Bereich „Website-Profil“ und gab dort eine Beschreibung der Website sowie weitere Kontaktdata an. Daraufhin überarbeitete die KI die Website. Bei der anschließenden Überprüfung der Veränderungen stellte er fest, dass die KI einen Sektor „Unsere Werte“ hinzugefügt hatte. Inhaltlich gefiel ihm dieser Abschnitt, insbesondere die Erwähnung des Themas Nachhaltigkeit. Allerdings fiel ihm auf, dass sich die KI manchmal inhaltlich „weit aus dem Fenster lehnt“, da Angaben wie „Tee aus kontrolliertem Anbau“ eigenständig generiert wurden, ohne dass der Ersteller solche Angaben gemacht hatte. Er merkte an, dass solche unbelegten Behauptungen rechtliche Probleme verursachen könnte. Hier wäre eine Überprüfung und Überarbeitung in einem

realen Szenario nochmal notwendig. In diesem Fall beließ er es jedoch dabei, da er auf seiner Website tatsächlich Bio-Tee im Angebot hat. Diese Beobachtung deckt sich mit der Behauptung von Ralf Otte (2021), der den unreflektierten Einsatz von KI durch Laien als eine Gefahr benennt. Da die Ergebnisse KI-basierter Systeme nicht zwangsläufig korrekt sind, ist eine kritische Überprüfung nötig. Diese Thematik wurde bereits zu Beginn dieser Arbeit thematisiert und findet an dieser Stelle eine erneute Bestätigung.

Als Nächstes wechselte er zum Bereich „Website-Struktur“ und ergänzte dort einen Online-Shop, eine Kontaktseite und eine „Über uns“-Seite. Da der Inhalt „Über uns“ bereits als Vorschlag vorhanden war, fügte er dies als weitere Seite hinzu. Anschließend überprüfte er die Inhalte unter „Website-Beschreibung“ und stellte fest, dass die vorhandenen Angaben noch zutreffend waren.

Im nächsten Schritt wollte er das „Website-Design“ anpassen. Er gab der KI die Anweisung, die Grundfarbe in helle Pastellfarben zu ändern. Die KI führte die Anweisungen aus, jedoch nicht exakt in der gewünschten Farbe, die er sich vorgestellt hatte. Insbesondere fehlten einzelne Farben für die verschiedenen Abschnitte. Deshalb gab er neue Anweisungen, um Abschnittsfarben hinzuzufügen. Doch auch diesmal war das Ergebnis unbefriedigend, da die KI erneut nicht die gewünschten Pastelltöne verwendete. Frustriert wiederholte er die Anweisung, woraufhin die KI die gesamte Website wieder in einheitliche Farben umwandelte.

Die Testperson wurde frustrierter und entschied sich dazu eine Version zurückzugehen, da diese am besten gepasst hatte. Anschließend gab er die Anweisung mehr Rundungen hinzuzufügen, jedoch setzte die KI dies nicht um. Er gab erneut Anweisungen, die er diesmal konkreter formulierte, in dem er die KI bat die blauen Elemente rundlicher zu gestalten. Die KI führte diese Anweisungen jedoch nicht aus. Die Testperson wurde in diesem Schritt zunehmend wütender, da die Anweisungen nicht wie gewünscht umgesetzt wurden. Er entschied sich daraufhin, zu der Version sechs von neun zurückzukehren, da diese ihm am besten gefallen hatte. Die Möglichkeit, zwischen verschiedenen Versionen zu wechseln, empfand er als positiv.

Trotzdem versuchte er erneut, über die vorgefertigte Anweisung von WIX „Farben neu generieren“ eine passende Farbe zu erhalten, jedoch waren die Änderungen minimal und ebenfalls nicht zufriedenstellend. Anschließend probierte er die vorgefertigte Anweisung „Design neu generieren“ aus. Daraufhin stürzte die Seite ab und musste neu geladen werden. Er wurde wieder zu der Übersichtsseite zurückgeworfen und musste viele Angaben erneut angeben.

Diesmal entschied er sich jedoch bei Art der Website für die Bezeichnung “Teestube” statt “Tee Online-shop”, da er hoffte ein besseres Ergebnis zu erzielen. Tatsächlich gefiel ihm das Design jetzt besser, doch ein richtiger Online-Shop ist weiterhin nicht erkennbar. Da er bereits sehr frustriert war und befürchtet, dass die Website bei erneuten Anweisungen bezüglich des Designs erneut abstürzen könnte, beließ er es dabei und klickte auf Website generieren.

Nach einer längeren Wartezeit wurde die Website von der KI fertiggestellt und er konnte sie sich ansehen.

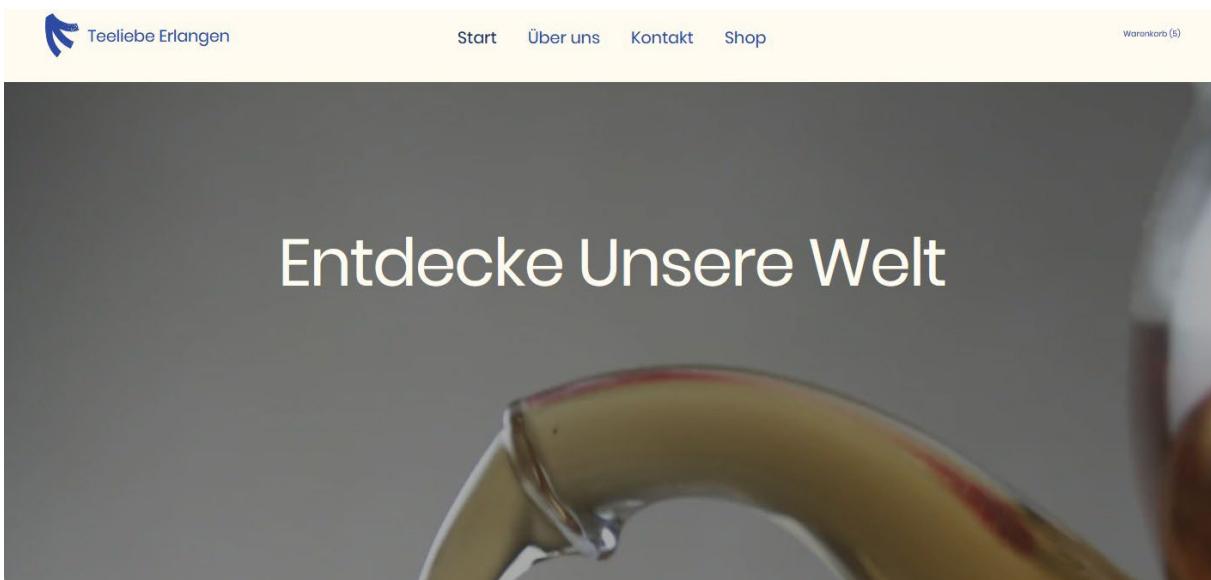


Abbildung 93. Finales Design der KI-Applikation

This image shows a section of the website titled 'Saisonale Favoriten' in large blue text. It displays three tea products: 'Bio Sencha' from 'Andere Regionen', 'Bio Spring' from 'China', and 'Bio China'. Each product has a small image, a name, and a price. A blue arrow icon is visible in the top right corner of the Bio China image.

Abbildung 94. Zweite Ansicht des finalen Designs der KI-Applikation

Er empfand die Website als „ganz okay“. Er vermutete, dass die Erstellung mit der KI insgesamt schneller ging, als wenn er es selbst mit einer Vorlage gemacht hätte. Vor allem die bereits generierten Texte sind besser als bei den Vorlagen. Allerdings müssten die Texte und die Farben für eine reale Website angepasst werden. Da die Erstellung der Produkte offenbar ebenfalls manuell erfolgen muss und dies ebenfalls nochmal viel Zeit kosten wird, zweifelte er an der tatsächlichen Zeitersparnis. Insbesondere, da es viele Vorlagen für Onlineshops gibt, mit denen sich ein ähnliches Ergebnis möglicherweise ebenso schnell hätte erzielen lassen.

Während sich die Testperson die Website weiter ansah, fiel ihm zudem auf, dass es nun doch einen Online-Shop gibt. In diesem befanden sich allerdings nur wenige Produkte mit nicht passenden Inhalten.

Daraufhin suchte er nach einer Möglichkeit, die Produkte mit der KI zu erstellen, fand jedoch weder eine passende Option noch hilfreiche Tutorials im Internet. Daher ging er davon aus, dass die Produkte selbst erstellt werden müssen. Hierbei wurde jedoch angemerkt, dass er sich an die vorgeschlagenen Einstellungen der KI halten soll, um das Ergebnis der KI erstellten Website nicht zu verfälschen. Die vorgeschlagenen Einstellungen der KI bestanden aus dem Hinzufügen eines Produktbildes, der Vergabe einer Kategorie, die Angabe von Produktdetails wie der Produktnamen, einem Banner, einer Beschreibung, die Preisgestaltung, die Produktoptionen, die aus der Angabe der verschiedenen Größen besteht und den Varianten, die den Preis pro Größe bestimmen.

Anschließend begann er mit der Erstellung des ersten Produkts, indem er die Auftragsbestätigung mit allen benötigten Produkten durchging. Bei der Erstellung der Produktbeschreibung fiel ihm auf, dass die KI einige Texte auf Englisch statt auf Deutsch generiert hatte. Dies verwirrte ihn und er musste entweder die Texte neu generieren lassen oder sie im Google Übersetzer ins Deutsche übertragen lassen.

Auch bei der Preisgestaltung gab es zunächst Probleme, da er verschiedene Preise je nach Größe angeben wollte, jedoch keine passenden Eingabeoptionen dafür fand. Er schaute sich die möglichen Eingaben, wie den Preis pro Einheit an, in der Hoffnung damit etwas verändern zu können, erzielte damit jedoch nicht das gewünschte Ergebnis. Schließlich gab er zunächst den Preis pro hundert Gramm an und scrollte weiter, um nach weiteren Einstellungen zu suchen. Schließlich stieß er auf die Möglichkeit, Produktoptionen zu definieren. Nach dem Lesen der Erklärung stellte er fest, dass er hier verschiedene Größen angeben kann. Allerdings gab es immer noch keine Möglichkeit, die Preise für jede Größe festzulegen. Verwirrt sah er sich weiter nach einer möglichen Einstellung um und wurde unter dem Punkt „Varianten verwalten“ fündig und konnte anschließend die Preise pro Größe anpassen.

Anschließend überprüfte er alle bisherigen Einstellungen, die er bisher getätigt hatte. Daraufhin entdeckte er die Option, Kategorien zu erstellen und eine bereits vordefinierte Kategorie „All Products“. Daraufhin legte er zusätzlich die Kategorie „grüner Tee“ an, damit es den Kunden möglich ist die Produkte besser zu unterscheiden. Er wollte zudem auch die Unterkategorie „Japan“ für eine weitere Unterscheidung ergänzen, fand jedoch keine passenden Einstellungen dafür. Stattdessen erstellte er versehentlich „Japan“ als eigenständige Kategorie auf derselben Ebene wie „Grüner Tee“. Da er keine Möglichkeit sah, dies zu korrigieren, ließ er es zunächst so stehen. Später fand er jedoch im rechten Menü unter „Produkte & Services“ eine Übersicht aller Kategorien, wo er diese anpassen und löschen konnte. Anschließend suchte er weiter nach einer Möglichkeit, die Produkte noch weiter zu unterscheiden und stieß dabei auf die Option „Banner“ unter den Produktdetails. Die KI hatte bisher nur das Wort „Banner“ eingetragen. Zunächst war er sich jedoch unsicher, wie sich diese Eingabe visuell auswirkt und schaute sich die Produktübersicht in der Vorschau der Website an. Dort sah er am oberen linken Rand des Produktbildes einen blau hinterlegten Text, mit der Inschrift „Banner“. Obwohl der Ersteller dies als nicht gut sichtbar empfand, nutzte er die Banner dennoch als weitere Unterscheidungsmöglichkeit.

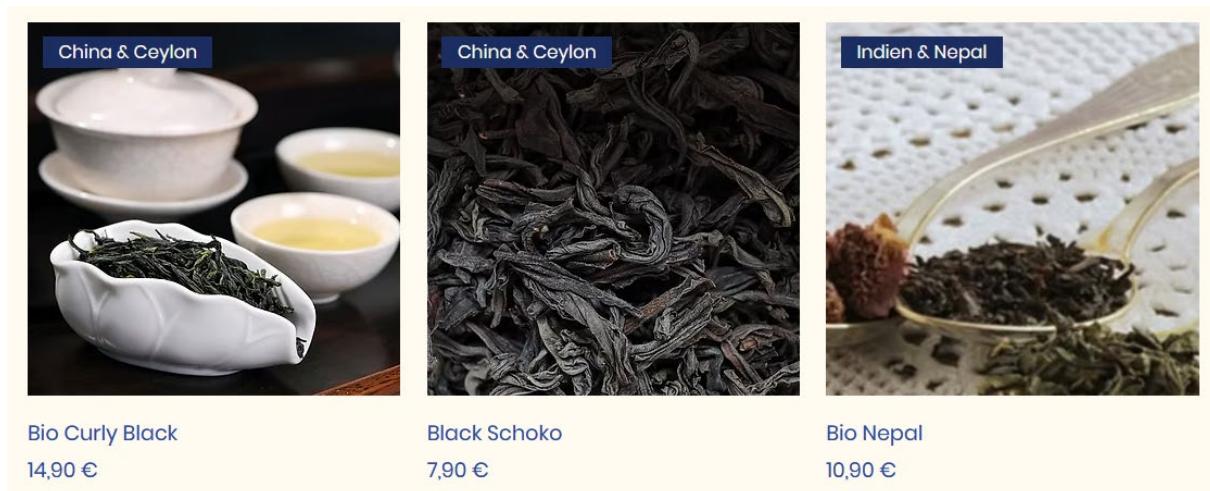


Abbildung 95. Banner zur weiteren Unterscheidung der Teesorten

Als nächstes wollte er das Header-Bild des Shops für „All Products“ sowie für die einzelnen selbst erstellten Kategorien ändern. Auch hier musste er zunächst eine Weile suchen bis er die Einstellung im rechten Menü unter „Produkte & Services“ und „Kategorien“ finden konnte. Unter diesem Menüpunkt konnte er den Kategorien ein Bild und eine Beschreibung hinzufügen. Zunächst war er unsicher, ob diese Änderungen erlaubt sind oder zu stark in das KI-Design eingreifen. Da die KI diese Einstellungen bereits für die Kategorie „All Products“ vorgenommen hatte, ging er jedoch davon aus, dass er dem gleichen Muster folgen darf. Also setzte er die Änderungen für alle elf Kategorien um, die er aus der Auftragsstellung übernommen hatte. Für jede dieser Kategorien mussten ebenfalls die passenden Produkte manuell erstellt werden. Dies stellte sich als sehr zeitintensiv heraus. Zwar konnte das erste Produkt dupliziert werden, doch jede Kopie musste entsprechend der neuen Angaben nachbearbeitet werden. Hier hätte sich der Ersteller mehr Unterstützung durch die KI gewünscht, da dieser Prozess insgesamt mehrere Stunden über mehrere Tage in Anspruch nahm.

The screenshot shows the final version of the online shop. At the top left is the logo "Teeliebe Erlangen". The top right features a navigation menu with links to "Start", "Über uns", "Kontakt", and "Shop". Below the menu is a search bar labeled "Suchen nach". To the right of the search bar is a large image of a tea infuser with dried green tea leaves. On the left side, there's a sidebar with a "Alle Produkte" section containing a list of categories like "Bio", "Früchtetee", "Genuss vor Ort", etc., and a "Filtern nach" section with checkboxes for "Produktart" (e.g., "Oolong Tee") and "Zubehör". The main content area is titled "Alle Produkte" and shows a grid of five recommended products under the heading "118 Produkte". Each product card includes a small image, the category (e.g., "Herhaftes", "Süßes"), and a brief description. A "Sortieren nach: Empfohlen" dropdown is visible at the top right of the product grid.

Abbildung 96. Finaler Online-Shop der KI-Applikation

Anschließend stellte der Ersteller fest, dass es im Erstellungsprozess keine Option gibt, Mitarbeiter anzugeben. Er überlegte, ob er diese Information unter „Über uns“ mithilfe der KI hinzufügen kann. Er durchsuchte zunächst den Website-Editor und die Einrichtungen, wo auch die Produkte verwaltet wurden, konnte jedoch keine entsprechende Einstellung finden. Also kehrte er in den Website-Editor zurück, öffnete die „Über uns“ Seite und versuchte, einen neuen Abschnitt mit der KI anzulegen. Da er sich nicht mehr erinnerte, wie dies funktionierte, musste er sich erneut das Tutorial ansehen. Anschließend konnte er einen Abschnitt mit der KI hinzufügen und musste lediglich einen Abschnittstyp auswählen, wobei er „Team“ und Keywords zu dem zu erstellenden Abschnitt angab. Dabei legte er fest, dass drei Mitarbeiter hinzugefügt werden sollten. Der erste Mitarbeiter stellt die Inhaberin Lena Tea selbst dar, Anna Li als Verkäuferin und Zuständige für den Bereich Marketing und Lena Johnsonn als Verkäuferin und Zuständige für den Bereich Verpackung. Nachdem die KI den Abschnitt ganz unten unter den Öffnungszeiten entwickelt hatte, bemerkte er, dass alle Texte auf Englisch sind. Er übersetzte die Überschrift manuell und lies die KI den Beschreibungstext neu generieren, in der Hoffnung, dass der Text diesmal auf Deutsch ist. Nachdem die KI den Text tatsächlich auf Deutsch generiert hatte, ersetzte er den Text auf Englisch links unter den Bildern mit den passenden Textstücken aus dem Text der KI rechts. Anschließend fiel ihm auf, dass die KI für Lena Johnsonn ein männliches Bild eingefügt hat. Er beschloss jedoch, dass dies keinen großen Unterschied macht und änderte den Namen Lena Johnsonn in Mark Johnsonn. Schließlich musste nur noch die Beschreibung der Tätigkeiten ins Deutsche übersetzt werden.

## Unser Team

Entdecken Sie die engagierten Menschen Teeliebe Erlangen. Lena Tea führt unser Team mit Kreativität und Hingabe. Anna Li ist für Marketing und Vertrieb zuständig und sorgt dafür, dass unsere Tees Sie erreichen. Mark Johnson kümmert sich um den Vertrieb und die Verpackung. sicherzustellen, dass Ihre Bestellungen einwandfrei sind.

<p>Lena Tea Inhaberin Lena Tea führt unser Team mit Kreativität und Hingabe.</p>	<p>Anna Li Verkäuferin &amp; Marketing Anna Li ist für Marketing und Vertrieb zuständig und sorgt dafür, dass unsere Tees Sie erreichen.</p>	<p>Mark Johnson Verkäuferin &amp; Verpackung Mark Johnson kümmert sich um den Vertrieb und die Verpackung, sicherzustellen, dass Ihre Bestellungen einwandfrei sind.</p>
--	--	--

Abbildung 97. Erstellter Abschnitt „Unser Team“

## Öffnungszeiten

---

Besuchen Sie uns	Montag - Freitag	9:00 - 18:00 Uhr
		Samstag

Abbildung 98. Öffnungszeiten der KI-generierten Applikation

In der Vorschau ging er die Website noch einmal durch und empfand die sie, als „gut genug“ mit den Änderungen, die ihm erlaubt waren. Weitere Verbesserungen waren mit der KI nicht mehr möglich. Außerhalb des Tests würde er jedoch noch einige Dinge ändern, die er mit der KI nicht zufriedenstellend umsetzen konnte. Zum einen müsste die Farbe der Schrift und zum anderen das Design allgemein geändert werden. Auch die Icons und das Logo hätte er gerne noch geändert. Da diese Anpassungen mit der KI schwierig, beziehungsweise unmöglich erschienen, beließ er die Website vorerst wie sie war.

### **3.2.3 Fazit**

Die Erstellung der Website mit dem KI-gestützten Website-Builder von WIX verlief insgesamt gut, brachte jedoch auch einige Herausforderungen mit sich. Während die Testperson den Erstellungsprozess und das Ausfüllen der vorgegebenen Angaben als hilfreich empfand, führten begrenzte Anpassungsmöglichkeiten, ungenaue Umsetzung von Anweisungen bezüglich des Designs und sprachliche Inkonsistenzen zu Frustration. Besonders die Gestaltung des Designs und die individuelle Anpassung von Farben, Produktkategorien und den Texten erwiesen sich als mühsam und nicht immer intuitiv.

Positiv wurde die Zeitersparnis durch vorgefertigte Inhalte und die einfache Navigation zwischen verschiedenen Versionen der Website bewertet. Dennoch zeigte sich, dass die KI-generierte Website in vielen Bereichen eine manuelle Nachbearbeitung erfordert, insbesondere bei den inhaltlichen Angaben und den Designentscheidungen. Trotz der KI-Unterstützung war der Zeitaufwand für die Erstellung der Website zudem höher als erwartet, da viele Details, wie die Entwicklung der Produkte, ohne der Hilfe der KI manuell generiert werden mussten. Insgesamt bietet die KI-basierte Erstellung nach dem Empfinden des Erstellers eine solide Grundlage, ersetzt jedoch nicht die Notwendigkeit menschlicher Nachkontrolle und Feinarbeit.

### **3.3 Testen der Lösungen gegen die Nutzungsanforderungen**

Mithilfe der in den folgenden Unterkapiteln beschriebenen Methoden wurde sowohl die manuell entwickelte Applikation als auch die KI-erstellte Applikation von einer Gruppe an Testpersonen getestet und bewertet (Burghardt et al., 2011).

#### **3.3.1 Usability Test**

Um ein Nutzerfeedback zu den beiden Applikationen zu erhalten, wurden moderierte Usability Tests mit insgesamt zweiundvierzig Testpersonen durchgeführt. Jede Testperson testete eine der beiden Applikationen, ohne im Voraus zu wissen, welche davon sie evaluieren würden, um eine Beeinflussung der Bewertung zu vermeiden. Der Datenerhebungsort war flexibel, musste jedoch eine ruhige Umgebung ohne Ablenkungen sein, in der sich die Testperson wohlfühlt. Die für den Test erforderliche Hardware umfasste einen Laptop, auf dem der Test durchgeführt wurde, einen externen Monitor für das Screen-Sharing zur besseren Beobachtung während des Tests sowie ein weiteres Gerät, wie ein iPad, zum Mitschreiben von Notizen für den Moderator. Die Testperson nahm dabei an einem Stuhl vor einem Tisch Platz. Um eine größere Anzahl an potenziellen Teilnehmern zu erreichen, konnten die Tests auch online über Zoom durchgeführt werden. Hierfür mussten die Teilnehmenden über einen Laptop mit eingeschalteter Kamera verfügen und das Screen-Sharing aktivieren. Nach der Anmeldung über Google Calender erhielten die Testpersonen eine Stunde vor dem Termin den Zugangslink und konnten zum vereinbarten Zeitpunkt dem Meeting beitreten.

Der Ablauf der Usability-Tests war standardisiert und wurde in sechs Abschnitte unterteilt.

Zunächst erfolgte die Vorbereitung der Tests. Hierfür wurden alle benötigten Materialien bereitgestellt, darunter ein Notizzettel für die Testperson, die Aufgaben ausgedruckt zur besseren Übersicht, die geöffnete Applikation auf dem Laptop, ein iPad für die Notizen während der Beobachtung sowie die vorbereitete Umfrage auf dem Laptop.

Daraufhin erfolgte eine Begrüßung der Testperson und eine kurze Einführung in das Projekt und das Ziel des Tests, um ein besseres Verständnis für den Kontext zu schaffen. Anschließend wurde der Testperson der Ablauf des Tests und die Methode des Usability-Tests vorgestellt. Besonders für unerfahrene Testpersonen war es wichtig, darauf hinzuweisen, dass es sich um einen Prototyp handelt, dessen Funktionen eingeschränkter sind als die einer voll entwickelten Website. Dies soll Verwirrung während der Tests vermeiden. Zudem wurde darauf hingewiesen, dass die Aufgaben nacheinander vom Moderator gestellt werden und auch schriftlich zum Mitlesen zur Verfügung stehen. Die Aufgaben gelten dabei als abgeschlossen, wenn die Testperson diese als abgeschlossen empfindet. Die Testpersonen wurden zudem gebeten, laut zu denken, also ihre Gedanken, Eindrücke und Handlungen während dem Testen der Applikation laut auszusprechen. Dies ermöglicht später eine genauere Nachverfolgung und Analyse der Interaktion. Zuletzt war es wichtig zu erwähnen, dass nicht die Testperson, sondern die Anwendung selbst getestet wird. Die Teilnehmenden konnten sich daher so viel Zeit nehmen, wie sie benötigten. Fehler oder Verwirrungen wurden nicht als Verschulden der Testperson, sondern als wichtige Erkenntnisse zur Optimierung der Applikation angesehen. Nachdem alle Fragen der Testperson geklärt waren, wurde der Test gestartet.

Anschließend begann der eigentliche Test. Wie bereits erwähnt, erhielt die Testperson die Aufgaben nacheinander. Währenddessen wurden, basierend auf den Äußerungen des Testers sowie den Beobachtungen des Moderators, Notizen gemacht. Geachtet wurde dabei auf die Emotionen, das Verhalten und die ausgesprochenen Gedanken und Gefühle hinsichtlich der Benutzerfreundlichkeit, der Benutzererfahrung oder dem Design. Zudem wurde die Interaktion mit der Applikation analysiert, insbesondere das Verständnis und der Umgang mit der Navigation und dem Interface sowie mögliche Schwierigkeiten bei der Bedienung. Fehler bei der Aufgabenlösung wurden ebenso notiert wie Erfolge, beispielsweise in Form von richtigen und schnellen Schritten in der Aufgabenlösung. Darüber hinaus wurden die Anzahl erfolgreich abgeschlossener und nicht abgeschlossener Aufgaben sowie die insgesamte Bearbeitungszeit erfasst. Diese Daten dienen später dem Vergleich der beiden Applikationen hinsichtlich der Benutzerfreundlichkeit und der Benutzererfahrung.

Anschließend füllte die Testperson anonym die Umfrage in Google Formulare aus.

Zum Abschluss wird dem Teilnehmer für seine Teilnahme an dem Test gedankt. Zudem wird das Gewinnspiel vorgestellt, bei dem als Dank jeweils zwei Gutscheine im Wert von zehn Euro verlost werden. Wenn die Person teilnehmen möchte, kann sie sich mit ihrem vollständigen Namen sowie einer E-Mail-Adresse in eine vorbereitete Excel-Liste eintragen.

### **3.3.2 Umfragen**

Um eine anonyme Bewertung der Applikationen zu erhalten, wurde eine Umfrage mit Google Formulare erstellt. Die Fragen basieren dabei hauptsächlich auf dem User Experience Questionnaire, kurz UEQ, einem schnellen und zuverlässigen Fragebogen zur Messung der User Experience interaktiver Produkte. Dabei wird die Benutzerfreundlichkeit anhand der Attraktivität, der Verständlichkeit und der Zuverlässigkeit abgefragt. Die Benutzererfahrung wird anhand der Bewertung, ob das Produkt spannend und motivierend ist sowie der Originalität und der Kreativität des Produktdesigns gemessen (Hinderks et al., n.d.). Der User Experience Questionnaire umfasst sechsundzwanzig Fragen, die in Form von semantischen Differenzialen gestaltet sind. Dabei bewerten die Befragten die Applikation auf einer Skala von eins bis sieben zwischen zwei gegensätzlichen Begriffen. Die siebenteilige Skala wird verwendet, um die Tendenz zur mittleren Bewertung zu verringern. Zudem werden die Begriffe in zufälliger Reihenfolge angeordnet, sodass bei der Hälfte der Fragen der positive Begriff zuerst steht und bei der anderen Hälfte der negative Begriff (Schrepp, 2023).

**Beispiel:**

Ich empfinde den Prototyp als...



Abbildung 99. Beispielfrage des UEQ

Diese Beispielfrage stammt aus der Umfrage zur Bewertung der manuell gestalteten Applikation. Der Bewerter kann diese auf einer Skala von eins bis sieben bewerten, wobei eins in diesem Beispiel die positivste Antwort, vier eine neutrale Antwort und sieben die negativste Antwort darstellt. Mit dieser Beurteilung sagt die Testperson aus, dass er den Prototyp eher attraktiv als unattraktiv empfindet. Dieses Beispiel wurde den Teilnehmern zu Beginn der Umfrage vorgestellt, um für ein besseres Verständnis zu sorgen.

Ich empfinde den Prototyp als... \*



Abbildung 100. Beispielfrage für die zufällige Reihenfolge der Begriffe des UEQ

Auch diese Beispielfrage stammt aus der Umfrage zur Bewertung der manuell gestalteten Applikation und veranschaulicht die zufällige Reihenfolge der Begriffe. In dieser Einschätzung kann die Testperson den Prototyp auf einer Skala von eins bis sieben bewerten, wobei eins in diesem Beispiel die negativste Antwort, vier eine neutrale Antwort und sieben die positivste Antwort darstellt. Mit dieser Beurteilung sagt der Bewerter aus, dass er den Prototyp eher unerfreulich als erfreulich empfindet.

Zusätzlich wurde die Umfrage um eigene Fragen erweitert, um die Bewertung der Navigation beziehungsweise des Zurechtfindens auf der Website, die Bewertung der Sortierung der Produkte, die Bewertung der Vollständigkeit der Inhalte und Funktionen und die Bewertung des Designs der Website noch genauer zu erfassen. Darüber hinaus kann die Testperson angeben, welche Inhalte und Funktionen ihr noch fehlen. Am Ende der Umfrage besteht zudem die Möglichkeit, freie Anmerkungen oder Verbesserungsvorschläge zu ergänzen.

Die Beispielfrage, die Fragen des User Experience Questionnaire und die eigenen Fragen sind dabei in einzelne Abschnitte unterteilt, um eine klare Struktur, eine bessere Übersichtlichkeit und eine leichtere Orientierung zu gewährleisten, indem die Fragen nach zusammenhängenden Gruppen sortiert sind.

## 4. Ergebnisse des Usability-Tests

In den folgenden Unterkapiteln werden auf die Ergebnisse des Usability-Tests mit Hilfe des Mann-Whitney-U-Tests und der deskriptiven Statistik eingegangen. Zudem werden die Mitschriften und Umfragen ausgewertet und der Unterschied der abgeschlossenen Aufgaben wird betrachtet.

### 4.1 Ergebnisse des Mann-Whitney-U-Test

Zur Überprüfung der statistischen Signifikanz wurde ein Mann-Whitney-U-Test durchgeführt, da es sich um zwei unabhängige Stichproben handelt und die Daten nicht normalverteilt sind. Im Folgenden werden die Ergebnisse des Tests vorgestellt. Dabei werden die statistisch signifikanten und die statistisch nicht signifikanten Datensätze getrennt betrachtet.

#### 4.1.1 Statistisch signifikante Datensätze

Im Folgenden werden die Datensätze dargestellt, bei denen der Unterschied zwischen der KI-generierten Applikation und der manuell erstellten Applikation in Bezug auf die abhängige Variable statistisch signifikant war, da der p-Wert kleiner als das zuvor festgelegte Signifikanzniveau von 0,05 ist. Die Nullhypothese  $H_0$  wird damit abgelehnt. Die Effektstärken zeigten zudem, dass der Unterschied auch praktisch bedeutsam ist.

$H_0$  = Es gibt keinen Unterschied zwischen den Gruppen in Bezug auf die abhängige Variable.

$H_1$  = Es gibt einen Unterschied zwischen den Gruppen in Bezug auf die abhängigen Variable.

Tabelle 2. Datensätze bei denen der Unterschied statistisch signifikant ist

Datensatz	U-Wert	p-Wert	Effektstärke
Übersichtlichkeit	U = 64	$p = 0,0002 < \alpha = 0,05$	$r = 0,63 \rightarrow$ großer Effekt
Unterstützend	U = 65,5	$p = 0,0002 < \alpha = 0,05$	$r = 0,62 \rightarrow$ großer Effekt
Verständlichkeit	U = 65	$p = 0,0002 < \alpha = 0,05$	$r = 0,62 \rightarrow$ großer Effekt
Effizienz	U = 88	$p = 0,0006 < \alpha = 0,05$	$r = 0,53 \rightarrow$ mittlerer Effekt
Gut	U = 97	$p = 0,0012 < \alpha = 0,05$	$r = 0,50 \rightarrow$ mittlerer Effekt
Erfreulich	U = 99,5	$p = 0,0018 < \alpha = 0,05$	$r = 0,48 \rightarrow$ mittlerer Effekt
Anziehend	U = 100	$p = 0,0018 < \alpha = 0,05$	$r = 0,48 \rightarrow$ mittlerer Effekt
Interessant	U = 99,5	$p = 0,002 < \alpha = 0,05$	$r = 0,48 \rightarrow$ mittlerer Effekt

Navigation	$U = 104,5$	$p = 0,002 < \alpha = 0,05$	$r = 0,48 \rightarrow$ mittlerer Effekt
Angenehm	$U = 110,5$	$p = 0,0046 < \alpha = 0,05$	$r = 0,44 \rightarrow$ mittlerer Effekt
Spannend	$U = 110,5$	$p = 0,005 < \alpha = 0,05$	$r = 0,43 \rightarrow$ mittlerer Effekt
Erwartungskonform	$U = 117$	$p = 0,0066 < \alpha = 0,05$	$r = 0,42 \rightarrow$ mittlerer Effekt
Sortierung	$U = 122$	$p = 0,01 < \alpha = 0,05$	$r = 0,40 \rightarrow$ mittlerer Effekt
Einfach	$U = 125$	$p = 0,012 < \alpha = 0,05$	$r = 0,39 \rightarrow$ mittlerer Effekt
Konservativ	$U = 123,5$	$p = 0,00124 < \alpha = 0,05$	$r = 0,39 \rightarrow$ mittlerer Effekt
Neuartig	$U = 126$	$p = 0,0136 < \alpha = 0,05$	$r = 0,38 \rightarrow$ mittlerer Effekt
Attraktiv	$U = 125,5$	$p = 0,0136 < \alpha = 0,05$	$r = 0,38 \rightarrow$ mittlerer Effekt
Pragmatisch	$U = 129$	$p = 0,0182 < \alpha = 0,05$	$r = 0,36 \rightarrow$ mittlerer Effekt
Sympathisch	$U = 130$	$p = 0,0198 < \alpha = 0,05$	$r = 0,36 \rightarrow$ mittlerer Effekt
Voraussagbar	$U = 132,5$	$p = 0,0198 < \alpha = 0,05$	$r = 0,36 \rightarrow$ mittlerer Effekt
Aufgeräumt	$U = 135$	$p = 0,0232 < \alpha = 0,05$	$r = 0,35 \rightarrow$ mittlerer Effekt
Originell	$U = 135,5$	$p = 0,03 < \alpha = 0,05$	$r = 0,34 \rightarrow$ mittlerer Effekt
Design	$U = 137$	$p = 0,032 < \alpha = 0,05$	$r = 0,33 \rightarrow$ mittlerer Effekt
Wert	$U = 140,5$	$p = 0,0394 < \alpha = 0,05$	$r = 0,32 \rightarrow$ mittlerer Effekt
Zeit	$U = 142$	$p = 0,048 < \alpha = 0,05$	$r = 0,3 \rightarrow$ mittlerer Effekt

#### 4.1.2 Statistisch nicht signifikante Datensätze

Im Folgenden werden die Datensätze dargestellt, bei denen der Unterschied zwischen der KI-generierten Applikation und der manuell erstellten Applikation in Bezug auf die abhängige Variable statistisch nicht signifikant war, da der p-Wert größer als das zuvor festgelegte Signifikanzniveau von 0,05 ist. Die Nullhypothese  $H_0$  wird damit beibehalten.

$H_0$  = Es gibt keinen Unterschied zwischen den Gruppen in Bezug auf die abhängige Variable.

$H_1$  = Es gibt einen Unterschied zwischen den Gruppen in Bezug auf die abhängigen Variable.

Tabelle 3. Datensätze bei denen der Unterschied statistisch nicht signifikant ist

Datensatz	U-Wert	p-Wert	Effektstärke
Schnelligkeit	$U= 146,5$	$p = 0,0524 > \alpha = 0,05$	$r = 0,3 \rightarrow$ mittlerer Effekt
Erlernbarkeit	$U= 153$	$p = 0,0734 > \alpha = 0,05$	$r = 0,28 \rightarrow$ kleiner Effekt
Aktivierend	$U= 152$	$p = 0,0802 > \alpha = 0,05$	$r = 0,27 \rightarrow$ kleiner Effekt

Kreativität	$U= 155,5$	$p = 0,097 > \alpha = 0,05$	$r = 0,26 \rightarrow$ kleiner Effekt
Sicherheit	$U= 172$	$p = 0,2006 > \alpha = 0,05$	$r = 0,2 \rightarrow$ kleiner Effekt

## 4.2 Ergebnisse der deskriptiven Statistik

Im Folgenden werden die Ergebnisse der deskriptiven Statistik für die Datensätze dargestellt, bei denen ein statistisch signifikanter Unterschied festgestellt werden konnte. Ermittelt wurde dabei der Mittelwert, der Median und der Interquartilsabstand.

Tabelle 4. Darstellung der Ergebnisse der deskriptiven Statistik

Datensatz	Mittelwert	Median	Interquartilsabstand
Übersichtlichkeit	KI = 3,95 Manuell = 1,62	KI = 4 Manuell = 1	KI = 3 Manuell = 1
Unterstützend	KI = 4,33 Manuell = 6,14	KI = 4 Manuell = 6	KI = 3 Manuell = 1
Verständlichkeit	KI = 4,43 Manuell = 6,43	KI = 5 Manuell = 7	KI = 3 Manuell = 1
Effizienz	KI = 4,19 Manuell = 6,14	KI = 5 Manuell = 6	KI = 3 Manuell = 1
Gut	KI = 3,43 Manuell = 1,66	KI = 3 Manuell = 1	KI = 3 Manuell = 1
Erfreulich	KI = 4,52 Manuell = 5,95	KI = 5 Manuell = 6	KI = 2 Manuell = 1
Anziehend	KI = 4,43 Manuell = 5,67	KI = 4 Manuell = 6	KI = 1 Manuell = 1
Interessant	KI = 3,95 Manuell = 5,52	KI = 4 Manuell = 6	KI = 2 Manuell = 2
Navigation	KI = 3,29 Manuell = 1,86	KI = 3 Manuell = 2	KI = 3 Manuell = 1
Angenehm	KI = 4,48 Manuell = 6	KI = 5 Manuell = 6	KI = 3 Manuell = 1
Spannend	KI = 3,52 Manuell = 4,95	KI = 3 Manuell = 5	KI = 2 Manuell = 2
Erwartungskonform	KI = 2,95 Manuell = 1,95	KI = 4 Manuell = 3	KI = 3 Manuell = 3
Sortierung	KI = 3,43 Manuell = 1,95	KI = 3 Manuell = 1	KI = 3 Manuell = 1
Einfach	KI = 5,29 Manuell = 6,9	KI = 6 Manuell = 6	KI = 1 Manuell = 1

Konservativ	KI = 3,1 Manuell = 4,24	KI = 3 Manuell = 4	KI = 1 Manuell = 2
Neuartig	KI = 2,62 Manuell = 3,81	KI = 2 Manuell = 4	KI = 1 Manuell = 3
Attraktiv	KI = 3,76 Manuell = 2,48	KI = 3 Manuell = 2	KI = 3 Manuell = 1
Pragmatisch	KI = 4,62 Manuell = 5,67	KI = 5 Manuell = 6	KI = 2 Manuell = 2
Sympathisch	KI = 3,67 Manuell = 2,33	KI = 3 Manuell = 2	KI = 3 Manuell = 2
Voraussagbar	KI = 5,24 Manuell = 6,19	KI = 6 Manuell = 6	KI = 1 Manuell = 1
Aufgeräumt	KI = 2,9 Manuell = 1,67	KI = 2 Manuell = 2	KI = 3 Manuell = 1
Originell	KI = 4,9 Manuell = 3,81	KI = 5 Manuell = 4	KI = 2 Manuell = 2
Design	KI = 3,24 Manuell = 2,33	KI = 3 Manuell = 2	KI = 2 Manuell = 2
Wert	KI = 3,62 Manuell = 2,33	KI = 4 Manuell = 2	KI = 3 Manuell = 2
Zeit	KI = 554s / 09:14 min Manuell = 437s / 07:17 min	KI = 537s / 08:57 min Manuell = 419s / 06:59 min	KI = 285s / 04:45 Manuell = 250s / 04:10 min

## 4.3 Auswertung der Mitschriften und der Umfragen

### 4.3.1 Feedback Capture Grid der KI-generierten Applikation

Zunächst wurde in Miro eine Vorlage für die Methode erstellt. Nach Abschluss der Usability-Tests wurden die Mitschriften und die Beobachtungen sowie Anmerkungen der Testpersonen ausgewertet. Zusätzlich wurden die Ergebnisse der begleitenden Umfragen ausgewertet und in einem separaten Feedback Capture Grid festgehalten.

#### Mitschrift

Positiv hervorgehoben wurde von allen einundzwanzig Testpersonen das einfache Auffinden der Filter, die gelungene Interaktion mit den Elementen sowie der übersichtliche und gewohnte Warenkorb, der zusätzlich eine Verlinkung zu PayPal enthält. Zudem berichteten dreizehn Testpersonen, dass die Applikation eine Lernkurve ermöglicht, wodurch das Auffinden weiterer Produkte nach der ersten Aufgabe deutlich erleichtert wurde. Acht Testpersonen lobten die zusätzliche Filtermöglichkeit Bio neben dem Filter Kräuter bei der Suche nach dem Bio-Kräutertee. Darüber hinaus wurden vereinzelt auch die Bilder

und Farben positiv bewertet, ebenso wie das angebotene Produktsortiment, die Auffindbarkeit der Banner und Produkte, das übersichtliche Warenkorb-Pop-up sowie die integrierte Gutscheinfunktion.

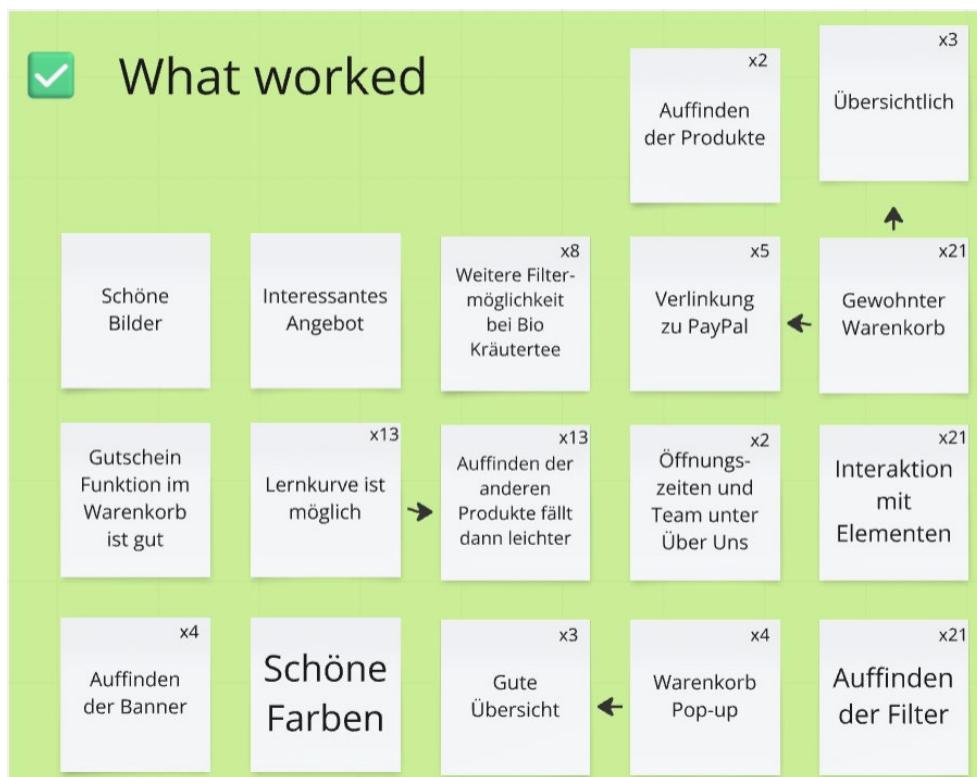


Abbildung 101. Feedback Capture Grid für die Mitschriften der KI-generierten Applikation

#### – What worked

Es ergaben sich jedoch auch einige Kritikpunkte wie die unklare Trennung von Matcha, Matcha-Zubehör und Zubehör. Auch die geringe Anzahl an Filtermöglichkeiten wurde häufig kritisiert, insbesondere das Fehlen einer Filterfunktion nach Herkunft, dass das gezielte Auffinden bestimmter Produkte erschwerte. Weiterhin war den Testpersonen die Unterscheidung der Filter meist nicht klar, da sie den Unterschied zwischen „Suchen nach“ und „Produktart“ nicht verstanden. Das Fehlen einer Suchfunktion, die das Finden der Produkte vereinfacht hätte, wurde ebenfalls bemängelt. Auch die Informationen zu den Öffnungszeiten und zum Team waren schwer auffindbar, insbesondere die Darstellung der Öffnungszeiten am Samstag führten bei vielen Testpersonen zu Unklarheiten. Des Weiteren wurde das Design der Website mehrmals negativ bewertet, insbesondere die Verwendung der blauen Schrift auf dem gelben Hintergrund, das Layout der Banner, das Logo sowie der Footer und dessen Inhalte. Der Warenkorb-Button wurde von mehreren Nutzern als zu klein empfunden. Zudem gaben zwei Testpersonen an, dass die Website zu viele visuelle Effekte hat. Der Online-Shop wurde als unübersichtlich wahrgenommen, als Grund wurde vor allem die Mischung der Tees nach Region und Sorte ohne zusätzlichen Filtermöglichkeiten angegeben. Zudem gab es zu wenige Informationen zu den Produkten in deren Unterseiten, vor allem das Fehlen der Information über die Herkunft des Tees viel negativ auf. Auch die Reihenfolge der Menüpunkte wurde als schlecht empfunden. Die Etagere war für viele Nutzer schwer auffindbar,

und die Verlaufsverlinkung wurde als verwirrend beschrieben. Schließlich wurde auch bemängelt, dass viele Buttons auf der Startseite nicht funktionierten.

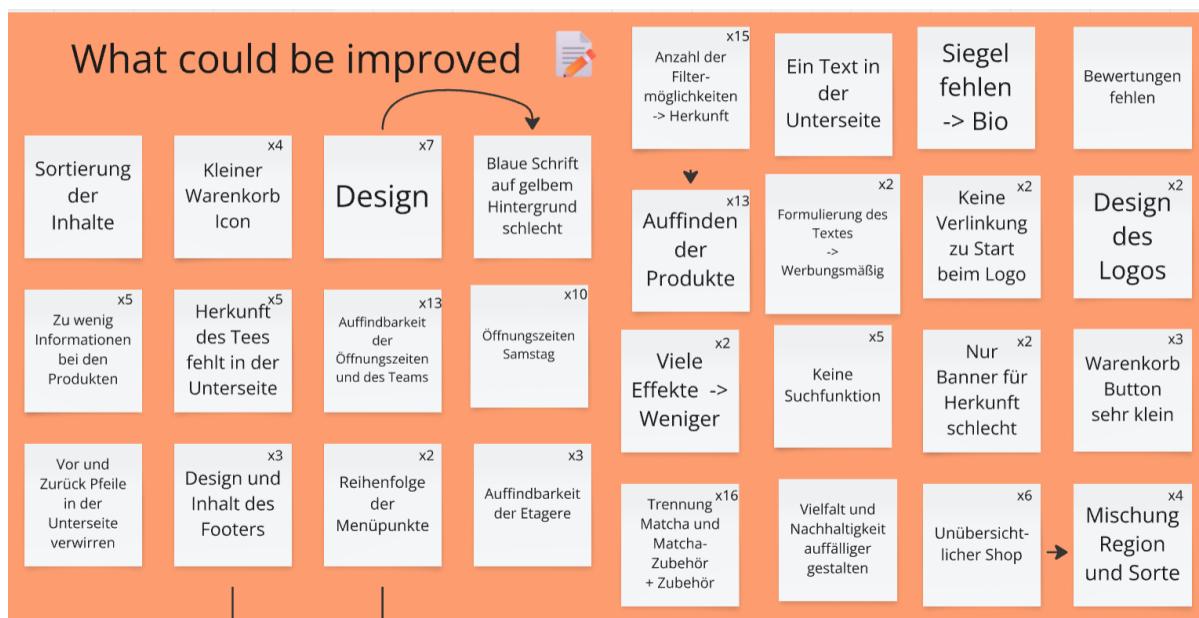


Abbildung 102. Feedback Capture Grid für die Mitschriften der KI-generierten Applikation

- What could be improved 1



Abbildung 103. Feedback Capture Grid für die Mitschriften der KI-generierten Applikation

## - What could be improved 2

Neben diesen Kritikpunkten hatten die Testpersonen jedoch auch einige Ideen wie die Website weiter verbessert werden könnte. So wurde vorgeschlagen, die Öffnungszeiten auch auf den Seiten „Start“, „Kontakt“ sowie im Footer darzustellen. Bezuglich der Reihenfolge der Menüpunkte wünschten sie sich, dass der Shop nicht so weit hinten platziert wird. Der Footer selbst sollte um rechtlich relevante Inhalte

wie Impressum und Datenschutz ergänzt und visuell klarer vom restlichen Seiteninhalt abgetrennt werden. Eine weitere Idee, war die Darstellung des Teams auch unter Start. Zudem sollten mehr Informationen zum Thema Tee angeboten werden. Das Einkaufs-Icon solle durch ein klassisches Einkaufswagen-Symbol ersetzt und die Abschnitte „Vielfalt“ und „Nachhaltigkeit“ auf der Startseite durch einen verschiedenen farbigen Hintergrund besser voneinander abgegrenzt werden. Auch für die Filter gab es Verbesserungsvorschläge. Es wurde zum einen vorgeschlagen, die Filter für Tee und Zubehör voneinander zu trennen und zum anderen auch Filter nach den Herkunfts ländern anzubieten, anstatt die Unterscheidung nur in Form von Banner darzustellen. Weiterhin wurde empfohlen, statt der verwirrenden Verlaufsanzeige das Menü in den Tee-Unterseiten am linken Rand zur besseren Orientierung zu ergänzen. Zuletzt wurde vorgeschlagen, die Textinhalte thematisch zu gliedern und in übersichtlicheren Stichpunkten darzustellen, um für eine bessere Lesbarkeit zu sorgen.

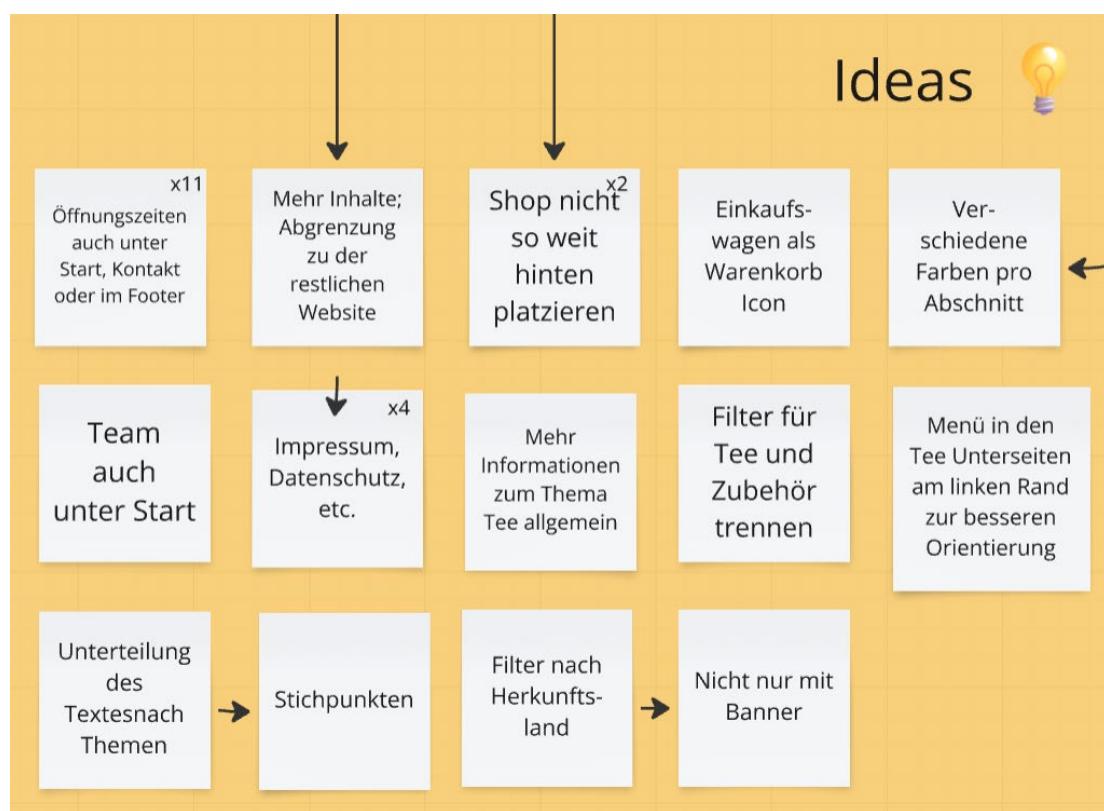


Abbildung 104. Feedback Capture Grid für die Mitschriften der KI-generierten Applikation

#### - Ideas

### Umfrage

Die Auswertung der Umfrage ergänzte die bereits aus den Mitschriften gewonnenen Erkenntnisse. Zum einen wurde die Navigation von sechs Personen als positiv hervorgehoben, insbesondere der bekannte Aufbau und die Übersichtlichkeit. Weitere sechs Personen bewerteten die vorhandenen Filterfunktionen als hilfreich, da diese passend zu den angebotenen Produkten nach Teesorten sortiert waren. Eine Person gab zudem an die Filter seien ausreichend. Auch das Design wurde von vier Personen positiv bewertet, wobei insbesondere die ansprechende, moderne, generische und zugleich schlichte

Gestaltung gelobt wurde. Weiterhin wurden die eingehaltenen Designregeln wie Abstand und Nähe als positiv erwähnt sowie die ins Gesamtkonzept passende Schriftgröße. Eine weitere Person gab zudem an, dass die die Website eigentlich selbsterklärend sei und für einfache Bestellungen ausreiche.

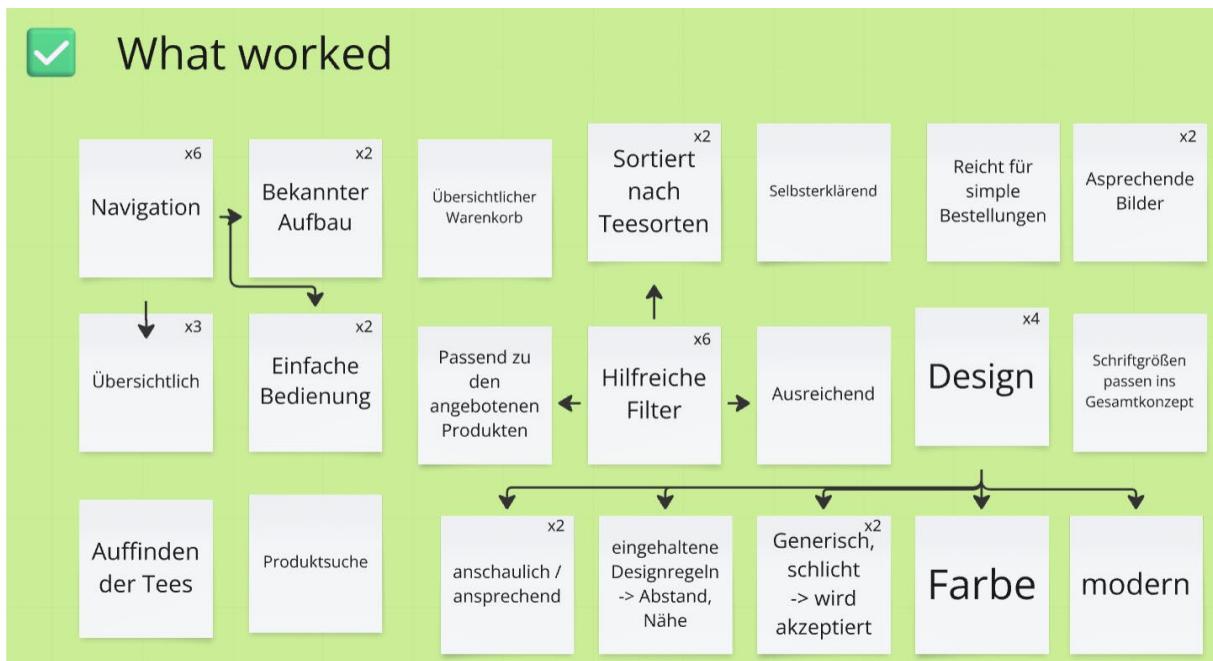


Abbildung 105. Feedback Capture Grid für die Umfrage der KI-generierten Applikation

#### - What worked

In der Umfrage ergaben sich auch zahlreiche Kritikpunkte, die bereits in den Mitschriften festgehalten worden waren. Es ergaben sich jedoch auch weitere Kritikpunkte. Drei Personen bemängelten, dass die Öffnungszeiten schwer und nicht an erwarteter Stelle auffindbar sind. Zwei weitere Personen kritisierten das Fehlen einer Bewertungsfunktion. Besonders deutlich fiel jedoch auch die Kritik am Design aus, da insgesamt dreizehn Personen dies kritisierten. Vor allem der Shop wurde als visuell wenig ansprechend und als „lustlos“ beschrieben. Allgemein wirkte das Design eher schlicht und langweilig auf die Nutzer. Fünf Personen bewerteten die Farben zudem als langweilig, unangenehm für die Augen und als nicht zusammenpassend. Weiterhin wurde bemängelt, dass es zu viele unsinnige Designelemente und Symbole ohne erkennbaren Sinn gibt. Auch die Texte sind zu lang, nicht einladend und haben eine komische und zu kleine Schriftart. Darüber hinaus wurde das Fehlen klarer Informationen zu den Versandkosten negativ hervorgehoben.



Abbildung 106. Feedback Capture Grid für die Umfrage der KI-generierten Applikation

### - What could be improved 1

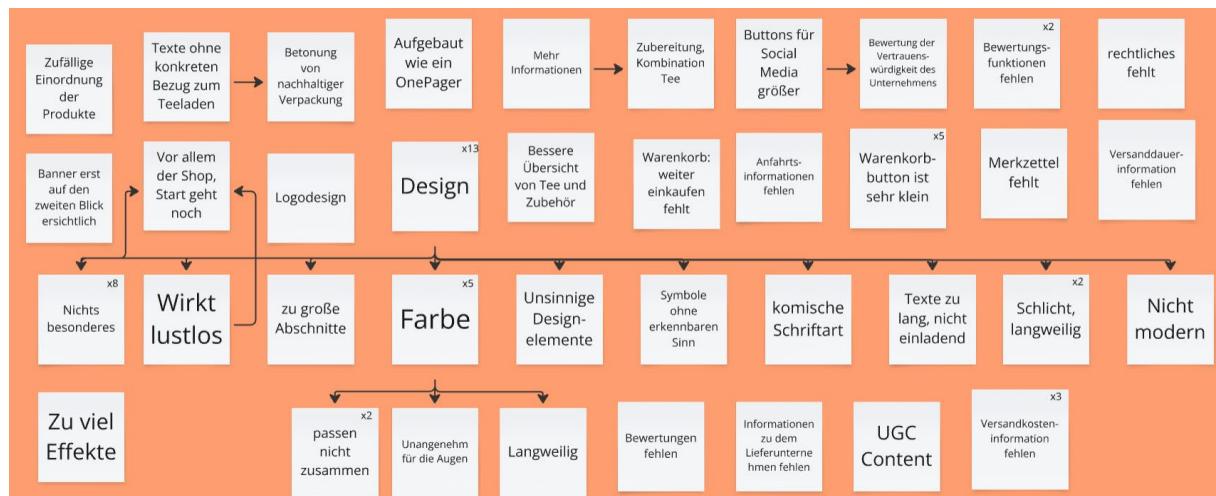


Abbildung 107. Feedback Capture Grid für die Umfrage der KI-generierten Applikation

### - What could be improved 2

In den Umfragen ergaben sich jedoch zusätzlich auch weitere Ideen zur Verbesserung der Website. Zum einen wünschten sich sechs Personen eine Suchfunktion, um gezielt nach bestimmten Schlagwörtern und speziellen Artikel suchen zu können. Außerdem wurde vorgeschlagen, einen eigenen Abschnitt für den Shop auf Home zu ergänzen. Weiterhin ergab sich die Idee einer Gramm-Angabe direkt bei der Produktvorschau. Zudem forderten mehrere Befragte mehr Informationen zur Zubereitung sowie zu möglichen Kombinationen der Tees. Auch sollte statt wie bisher hauptsächlich die Verpackung zu betonen, stärker auf die Qualität des Tees eingegangen werden. Eine weitere Idee stellte die Integration eines KI-gestützten Chatbots dar, der Fragen der Kunden beantworten könnte. Zuletzt wurde vorgeschlagen grüne, braune oder beige Farbtöne zu verwenden.



Abbildung 108. Feedback Capture Grid für die Umfrage der KI-generierten Applikation

- Ideas

### 4.3.2 Feedback Capture Grid der manuell gestalteten Applikation

Zunächst wurde in Miro eine Vorlage für die Methode erstellt. Anschließend wurde nach Abschluss der Usability-Tests die Mitschriften und die Beobachtungen sowie Anmerkungen der Testpersonen ausgewertet. Auch die Umfrageergebnisse wurden analysiert und in einem separaten Feedback Capture Grid festgehalten.

#### Mitschrift

Alle einundzwanzig Testpersonen konnten sicher mit den Elementen der Website interagieren. Sie fanden die Filter problemlos, wendeten diese an und konnten dadurch auch schnell und einfach alle Tees auffinden. Zwanzig Personen beschrieben die Filter zudem als hilfreich. Auch das Abschließen der gestellten Aufgaben gelang zwanzig Nutzern ohne großartige Probleme. Weitere neunzehn Personen bewerteten den Warenkorb als übersichtlich und gewohnt, da er in einzelne Schritte aufgeteilt ist und eine klare Struktur hat. Auch der Warenkorb-Pop-up wurde von mehreren Teilnehmern als positiv bewertet. Das Auffinden der Etagere viel neunzehn Personen und das Auffinden des Matcha-Sets sechzehn Personen leicht. Zehn Nutzer bewerteten die angebotenen Produktinformationen als interessant, insbesondere zur Herkunft und den verwendeten Qualitätssiegeln. Weitere acht Personen lobten die übersichtlich gestalteten Produktseiten und betonten, dass die Produkte auch ohne den Einsatz von Filtern leicht auffindbar sind. Das Design der Website wurde von vier Personen positiv bewertet, insbesondere auf Grund der schönen Farbauswahl. Auch die Verwendung bereits bekannter Funktionen, wie die Verlinkung des Logos zur Startseite oder das Anbieten von Bewertungen und den dazu passenden Filtern, wurde positiv aufgenommen. Fünf Personen gaben an, dass sie durch einen Lerneffekt noch schneller mit der Website interagieren konnten. Darüber hinaus wurden die Gutscheinfunktion, der Aufbau des Menüs, die Vergleichsliste, der Merkzettel sowie die Strukturierung der Startseite in einzelne Abschnitte als weitere positive Aspekte genannt.



Abbildung 109. Feedback Capture Grid für die Mitschrift der manuell gestalteten Applikation

- What worked 1



Abbildung 110. Feedback Capture Grid für die Mitschrift der manuell gestalteten Applikation

#### - What worked 2

Es ergaben sich jedoch auch Kritikpunkte. Fünf Testpersonen hatten Schwierigkeiten, das Matcha-Set zu finden, da die Unterscheidung der Filter Zubehör und Matcha als verwirrend empfunden wurde. Bei der Beendung des Bestellprozesses hatten drei Nutzer Probleme, da sie es als ungewohnt empfanden, vor Abschluss der Bestellung nicht direkt auf Bezahlseiten wie PayPal weitergeleitet zu werden. Zwei weitere Personen hatten Schwierigkeiten, die Etagere zu finden. Weitere vereinzelte Kritikpunkte stellten das Auffinden des Teams, die zu kleinen Buttons, das Buttondesign für die Merk- und Vergleichsliste, der zu viele Text und die nicht ansprechende grüne Farbe dar.

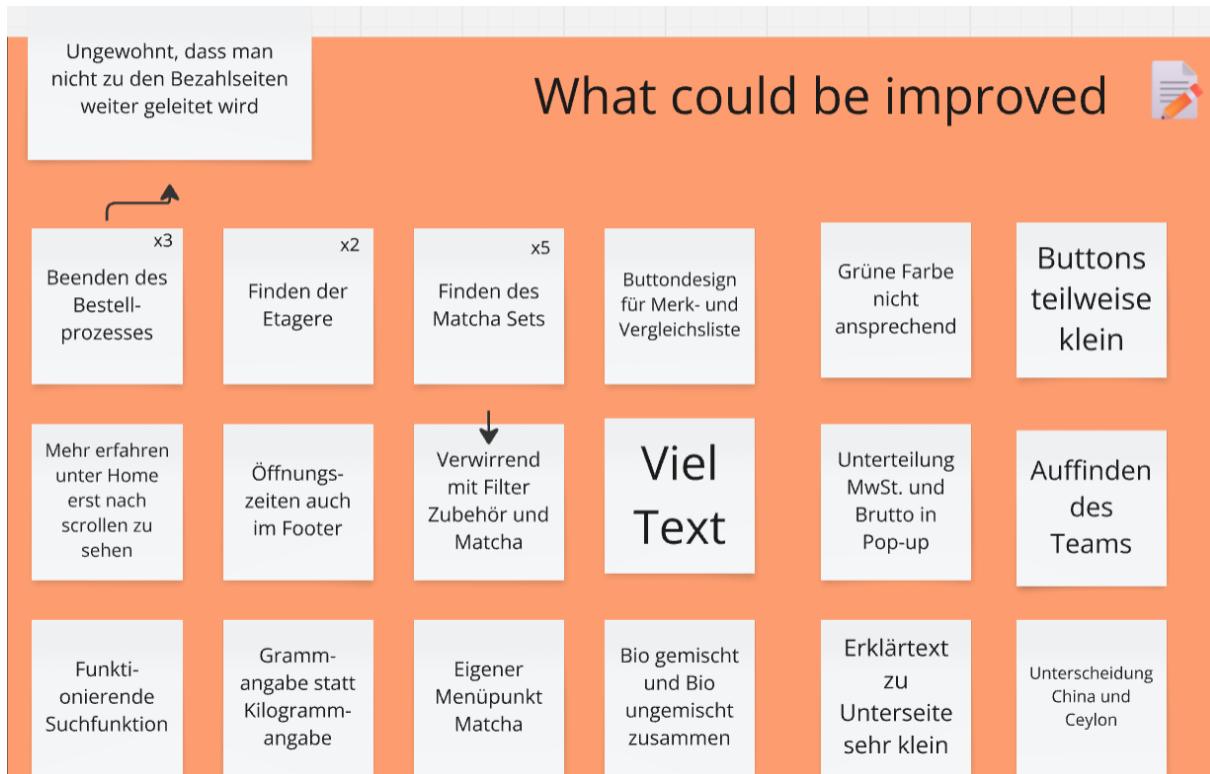


Abbildung 111. Feedback Capture Grid für die Mitschrift der manuell gestalteten Applikation

### - What could be improved 1

Mehrere Testpersonen hatten jedoch auch Ideen, wie die Applikation weiter verbessert werden könnte. Zum einen könnte ein Kundenlogin ergänzt werden, um persönliche Daten für einen erneuten Einkauf zu speichern und den Kaufprozess zu beschleunigen. Zum anderen könnten die aktuellen Filter durch Menüpunkte mit eigenen Unterseiten ersetzt werden. Auch der Bio-Filter könnte verbessert werden, indem er als Eingrenzung der Suche wie der Bestseller-Filter funktioniert. Auch die Einsortierung der Etagere könnte überdacht werden. Des Weiteren wurde vorgeschlagen, die Öffnungszeiten und Informationen über den Laden auch in den Bereich „Über den Laden“ zu integrieren sowie vermehrt Stichpunkte statt Fließtexte zu verwenden, um die Lesbarkeit zu verbessern.



Abbildung 112. Feedback Capture Grid für die Mitschrift der manuell gestalteten Applikation

#### - Ideas

Im Verlauf der Tests kamen auch zwei Fragen auf. Ein Nutzer fragte sich, was unter dem Filter „andere Regionen“ zu erwarten ist. Eine andere äußerte Unverständnis über die Höhe der Versandkosten bei Hermes und wünschte sich eine entsprechende Erklärung.

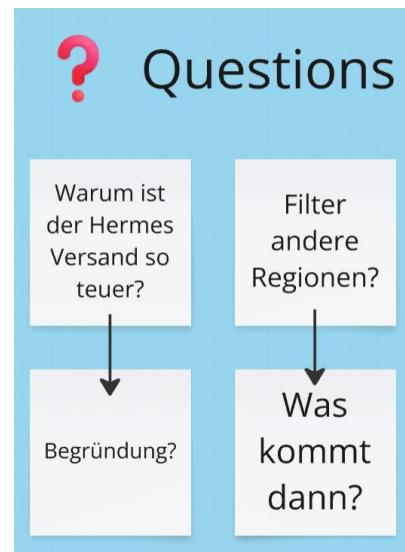


Abbildung 113. Feedback Capture Grid für die Mitschrift der manuell gestalteten Applikation

#### - Questions

## Umfrage

Die Auswertung der Umfrage ergänzte die bereits aus den Mitschriften gewonnenen Erkenntnisse. Zum einen wurde von zwanzig Nutzern die Navigation als positiv bewertet, da diese gut sortiert ist, sinnvoll und gewohnt aufgebaut ist und dank der Übersichtlichkeit, ein schnelles und einfaches Zurechtfinden ermöglicht. Zudem beschrieben die Navigationspunkte die Inhalte gut. Fünf Personen gaben an, die Produkte an erwarteten Orten gefunden zu haben und bewerteten die Strukturierung und Sortierung der Website positiv. Zwei weitere Testpersonen lobten die gute Übersicht durch die Über- und Unterkategorien, die von allgemein zu spezifisch gingen. Auch das Design der Website wurde von insgesamt zwölf Testpersonen als positiv bewertet. Die angenehme Farbwahl, die aufgeräumte und übersichtliche Gestaltung sowie das moderne, jedoch eher konservative Design wurden als unterstützend für ein schnelles Auffinden von Informationen und Produkten angesehen.

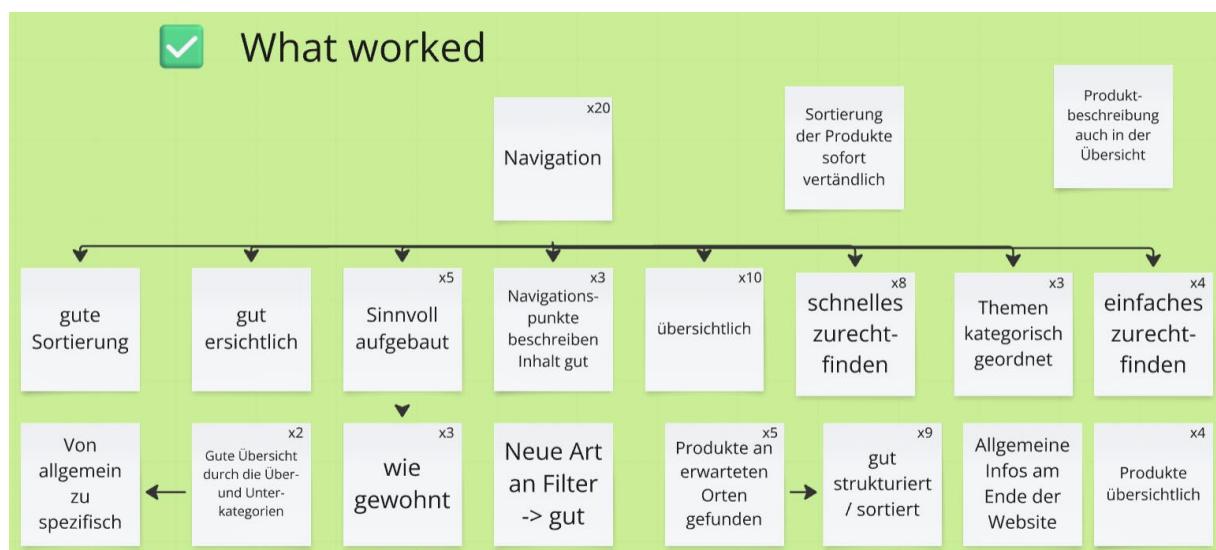


Abbildung 114. Feedback Capture Grid für die Umfrage der manuell gestalteten Applikation

- What worked 1

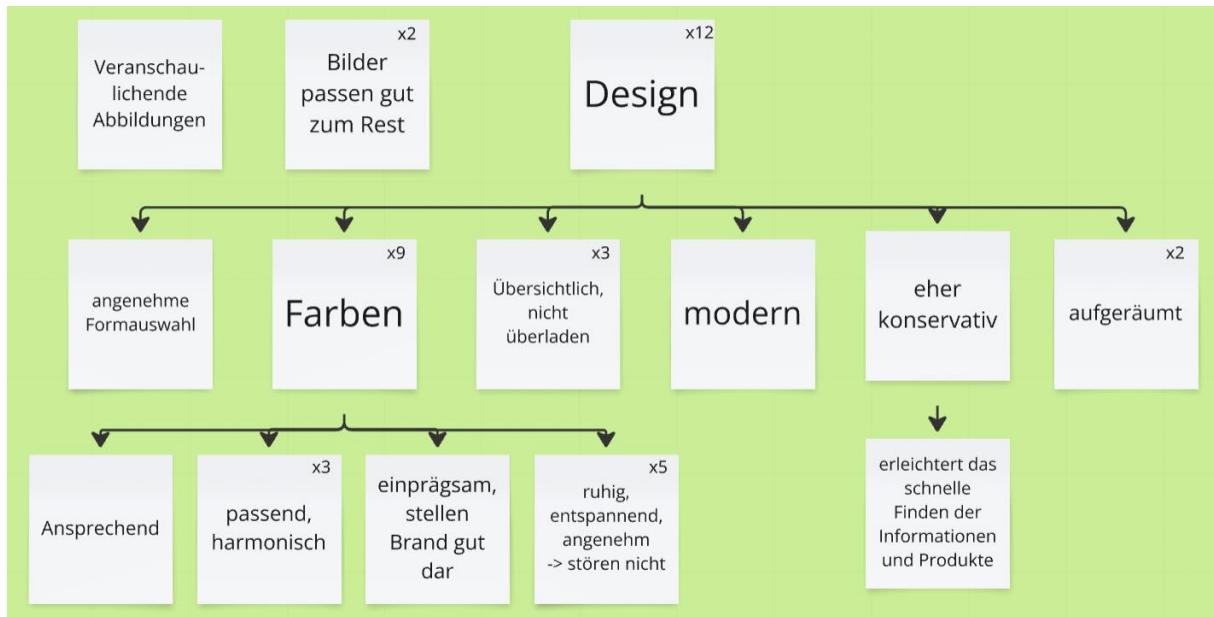


Abbildung 115. Feedback Capture Grid für die Umfrage der manuell gestalteten Applikation

#### - What worked 2

Gleichzeitig ergaben sich im Rahmen der Umfrage auch weitere kritische Anmerkungen. Neun Personen gaben an, dass das Design überarbeitet werden sollte, insbesondere hinsichtlich einer lebendigeren Farbgestaltung für mehr „Pepp“ und Tiefe. Hierbei wurde vor allem angemerkt, dass die Hintergrundfarbe in eine auffälligere Farbe geändert werden sollte. Zwei weitere Nutzer beschrieben die Website als lieblos, eintönig und nicht innovativ oder neuartig. Kritisiert wurde weiterhin die nicht ausreichende Hervorhebung des Logos sowie die nicht ansprechende Gestaltung der Buttons im Kaufprozess. Darüber hinaus wurde die Startseite als zu textlastig wahrgenommen.

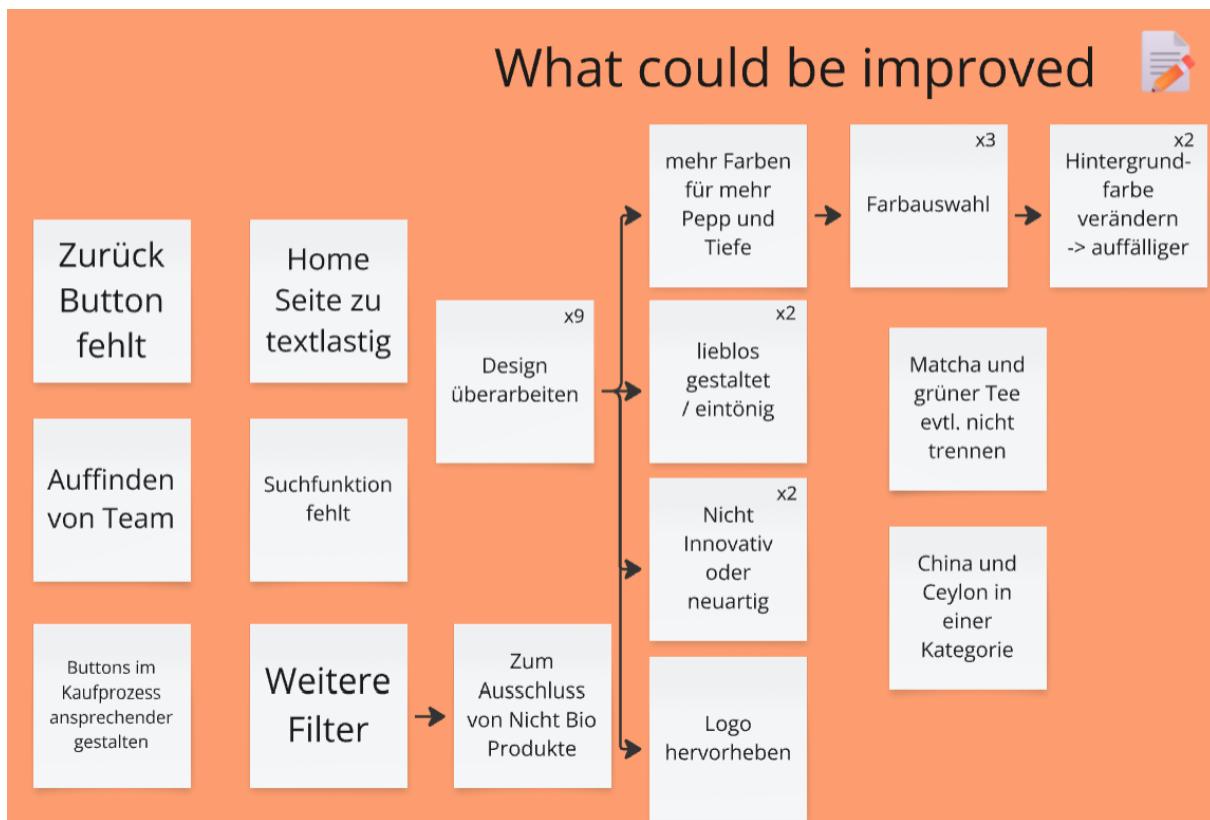


Abbildung 116. Feedback Capture Grid für die Umfrage der manuell gestalteten Applikation

#### - What could be improved

Es ergaben sich auch weitere Ideen, die aus den Mitschriften ergänzen. Zum einen könnte ein großes Bild unter der Menüleiste in Home eingefügt werden. Zudem wurde vorgeschlagen, Produktkombinationen anzubieten sowie kurze Informationsvideos über den Anbau und die Ernte des Tees bereitzustellen. Weitere Ideen wie das Anbieten von Teerezepten, Grammangaben der Tees und eine Rückmeldung nach der Bestellung sind in der Website bereits enthalten, wurden jedoch von den Testpersonen, die die Vorschläge angaben, nicht wahrgenommen.



Abbildung 117. Feedback Capture Grid für die Umfrage der manuell gestalteten Applikation

#### - Ideas

Eine Testperson hatte auch die Frage gestellt, warum zwischen den einzelnen Kategorien unterschieden wird, ohne näher darauf einzugehen.

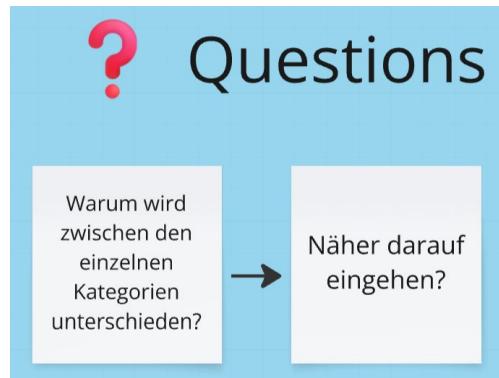


Abbildung 118. Feedback Capture Grid für die Umfrage der manuell gestalteten Applikation

#### - Questions

### 4.3.3 Emotionen während der Tests der KI-generierten Applikation

Während der Tests traten sowohl positive als auch negative Emotionen bei den Testpersonen auf. Insgesamt empfanden fünfzehn Testpersonen Freude bei der Nutzung der KI-generierten Applikation. Hauptsächlich auf Grund des gewohnten Warenkorbes, des Lerneffektes und der zusätzlichen Filtermöglichkeit Bio. Neun weitere Personen wiederum empfanden Erleichterung wegen dem gewohnten Warenkorb, der weiteren Filtermöglichkeit Bio und das schnelle Auffinden von Produkten durch die Banner. Eine Person zeigte zudem Interesse an den angebotenen Produkten sowie an der simplen und einfachen Website.

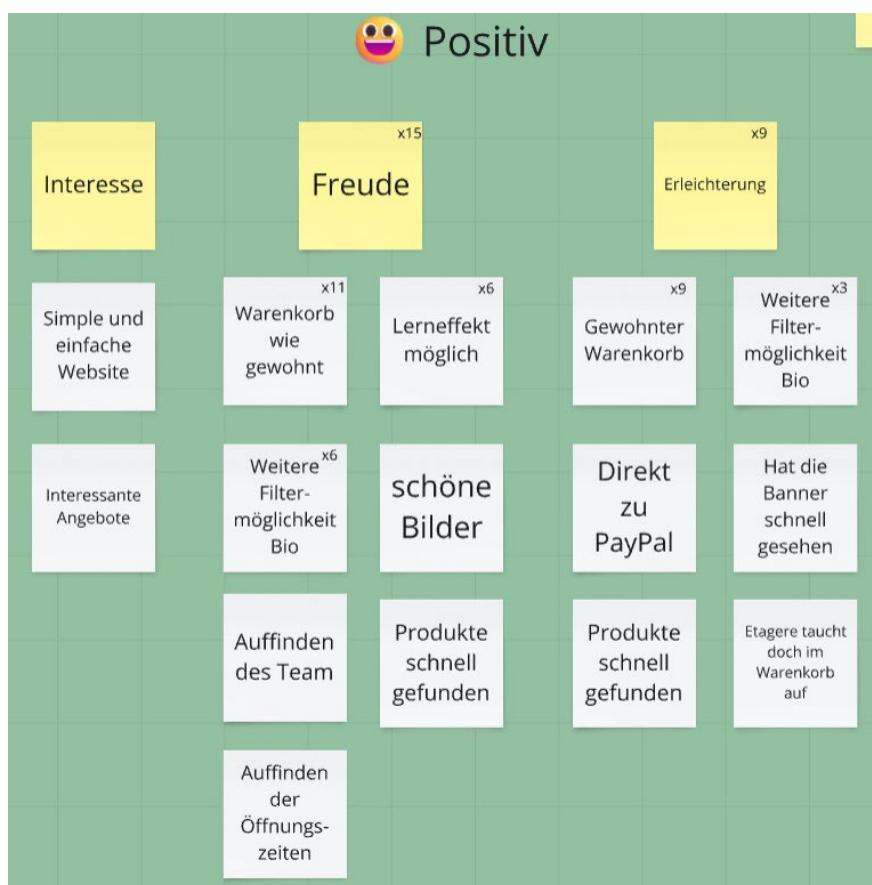


Abbildung 119. Positive Emotionen während des Nutzens der KI-generierten Applikation

Neben den positiven Reaktionen wurden jedoch auch zahlreiche negative Emotionen geäußert. Dabei empfanden achtzehn Personen Verwirrung, als es keine weiteren Filter gab, bei der unklaren Unterscheidung von Matcha und Matcha-Zubehör, dem Auffinden der Öffnungszeiten und des Teams sowie bei den Bannern. Sechzehn Testpersonen empfanden Entsetzen, als sie die Suche kaum mithilfe von Filter eingrenzen konnten, unter dem Menüpunkt Zubehör sowie allgemein sehr viel suchen mussten, den vollen Shop und den kleinen Warenkorb gesehen haben. Weitere vier Personen empfanden Wut, da ihnen weitere Filtermöglichkeiten für die gezielte Suche der Tees fehlten und sie dadurch nicht den gewünschten Tee fanden. Einige Testpersonen empfanden jedoch auch Belustigung, als sie den kleinen Warenkorb sahen oder bei dem Eindruck, dass bei der Gestaltung offenbar keine bewussten Designentscheidungen getroffen wurden. Auch Langeweile wurde empfunden, da die Website weder

spannende Inhalte noch weiterführende Informationen zu den Produkten anbot. Zuletzt kam auch Enttäuschung auf, als es keine weiteren Filter gab sowie weitere Informationen zu den Tees.



Abbildung 120. Negative Emotionen während des Nutzens der KI-generierten Applikation



Abbildung 121. Weitere Negative Emotionen während des Nutzens der KI-generierten Applikation

Eine Testperson hatte jedoch auch eher neutrale Emotionen während der Nutzung der Applikation. Das heißt er empfand weder starke Freude noch Frustration oder Wut. Er war eher kritisch und misstrauisch gegenüber der Website.

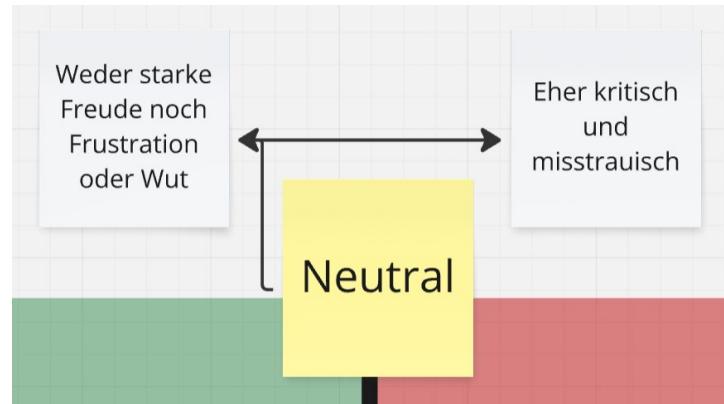


Abbildung 122. Neutrale Emotionen während des Nutzens der KI-generierten Applikation

#### 4.3.4 Emotionen während der Tests der manuell gestalteten Applikation

Während der Tests der manuell gestalteten Applikation traten sowohl positive als auch negative Emotionen auf. Insgesamt empfanden einundzwanzig Testpersonen Freude während der Ausführung der Aufgaben. Besonders beim schnellen und einfachen Auffinden der Produkte, unterstützt durch die praktischen Filter. Weitere Gründe für die empfundene Freude waren die interessanten Informationen, vor allem über die Herkunft des Tees, sowie der übersichtliche Aufbau der Website. Ebenfalls einundzwanzig Personen empfanden Ruhe, als sie die Bestellungen tätigten, da sie alle Aufgaben ohne Probleme abschließen konnten. Sieben Testpersonen zeigten Interesse, als sie sich die Produktinformationen, die Bewertungen und die Website insgesamt ansahen.

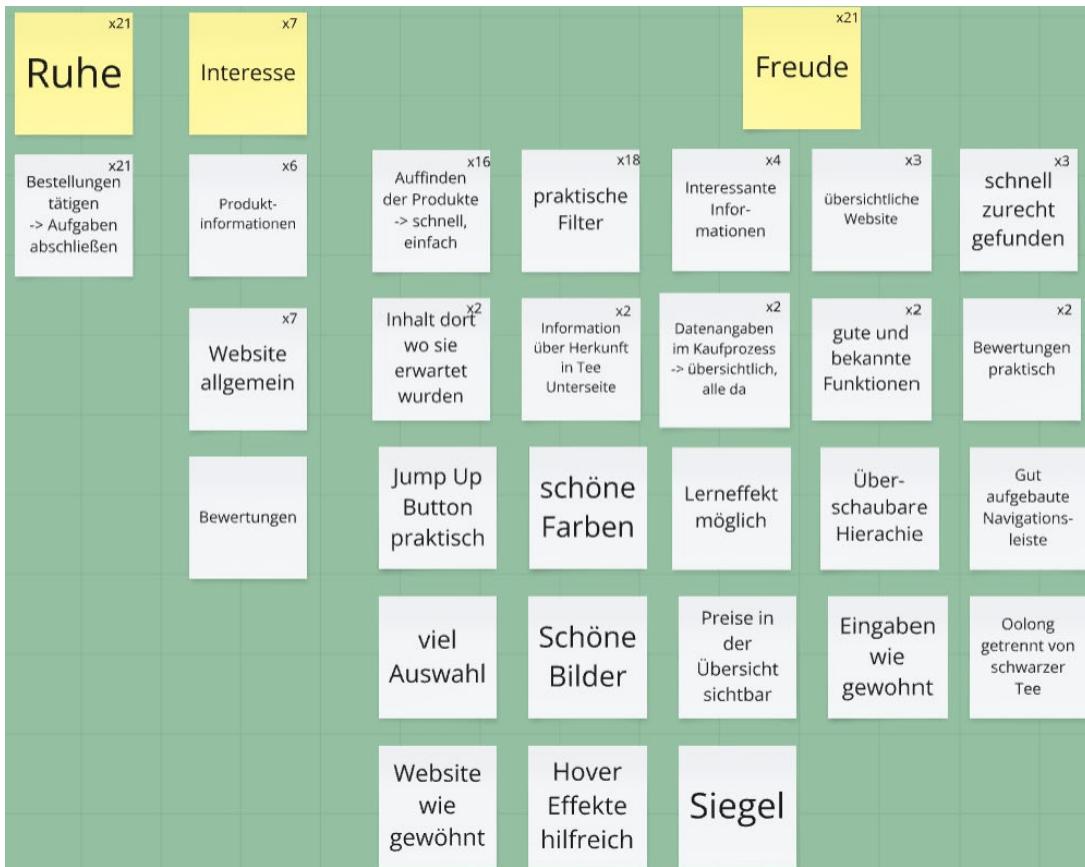


Abbildung 123. Positive Emotionen während des Nutzens der manuell gestalteten Applikation

Den positiven Eindrücken standen jedoch auch negative Emotionen gegenüber. Sieben Personen fühlten sich verwirrt, unter anderem weil zwei von ihnen sich nicht sicher waren, wann der Bestellprozess abgeschlossen war. Weiterhin ergaben sich vereinzelt Schwierigkeiten beim Auffinden des Teams, des Matchas und der Etagere. Eine Testperson empfand Enttäuschung, als es keinen Kundenlogin gab. Eine weitere Person war genervt von dem langen Fließtext ohne Stichpunkte, dem Fehlen genauer Informationen zu den Inhaltsstoffen, der fehlenden Grammangabe in der Produktübersicht und die Zusammenfassung der gemischten und ungemischten Bio-Kräutertees.



Abbildung 124. Negative Emotionen während des Nutzens der manuell gestalteten Applikation

#### 4.3.5 Verhalten während der Tests der KI-generierten Applikation

Während der Tests konnte sowohl positives als auch negatives Verhalten bei den Testpersonen beobachtet werden. Siebzehn Personen verhielten sich zunächst sicher während der Nutzung der Applikation, insbesondere beim Anwenden der vorhandenen Filter, bei der Interaktion mit dem bekannten Warenkorb sowie bei dem Auffinden der Öffnungszeiten, des Teams und der Produkte. Dieses Verhalten zeigte sich auch, als ein Lerneffekt nach der ersten Aufgabe einen sicheren Abschluss der weiteren Aufgaben ermöglichte.



Abbildung 125. Positives Verhalten während des Nutzens der KI-generierten Applikation

Neben den positiven Beobachtungen zeigten auch viele Testpersonen negatives Verhalten. Zwanzig Personen zeigten Unsicherheiten, vor allem beim Auffinden des Matcha-Sets, der Öffnungszeiten, der Etagere und der verschiedenen Teesorten. Vierzehn Testpersonen reagierten zudem genervt, als sie keine weiteren Filter fanden und es ihnen dadurch schwerfiel, bestimmte Produkte wie den Tee, das Matcha-Set oder die Etagere zu finden. Auch das Fehlen einer Suchfunktion, ein zu kleines Warenkorb-Icon sowie das Fehlen einer Verlinkung des Logos zur Startseite sorgte für genervte Reaktionen. Viele zögerten bei dem Suchen nach den Öffnungszeiten, dem Team, den Teesorten und dem Matcha-Set. Dieses Verhalten zeigte sich auch bei der Erkennung der Herkunft der Tees in den Unterseiten und dem weiteren Vorgehen, als sie keine weiteren Filter fanden. Drei Personen wirkten besonders verzweifelt, als sie versuchten das Team, die Öffnungszeiten oder das Matcha-Set zu finden. Eine Testperson erschrak beim Anblick der großen Anzahl an Produkten im Shop.



Abbildung 126. Negatives Verhalten während des Nutzens der KI-generierten Applikation



Abbildung 127. Weiteres negatives Verhalten während des Nutzens der KI-generierten Applikation

#### 4.3.6 Verhalten während der Tests der manuell gestalteten Applikation

Auch bei der manuell gestalteten Applikation zeigte sich sowohl positives als auch negatives Verhalten. Einundzwanzig Testpersonen verhielten sich sicher, insbesondere bei der Interaktion mit dem Interface und den Abschluss der Aufgaben, vor allem beim Auffinden der Produkte wie die Tees oder die Etagere. Eine Person zeigte sich erleichtert, als sie schließlich verstand, wann der Bestellprozess abgeschlossen ist.

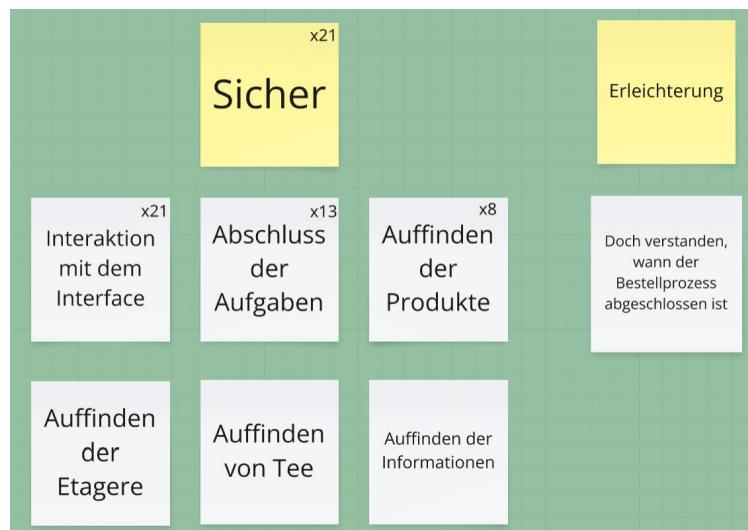


Abbildung 128. Positives Verhalten während des Nutzens der manuell gestalteten Applikation

Jedoch zeigten vereinzelte Testpersonen während der Tests auch negatives Verhalten. Vier Personen zögerten während des Abschlusses des Bestellprozesses, da sie nicht verstanden, wann der Prozess

abgeschlossen ist. Vier weitere Personen zeigten Unsicherheiten, insbesondere beim ersten Anwenden der Filter, da die Herkunft der Tees nicht direkt aus den Produktnamen ersichtlich war und sie befürchteten, dass die Filterfunktion nicht funktioniert haben könnte. Auch beim Auffinden der Etagere, des Matcha-Sets, der Öffnungszeiten und des Teams entstanden vereinzelt Unsicherheiten. Zudem verhielten sich vier Testpersonen nachdenklich, da sie unsicher waren, unter welchem Filter sie die Etagere oder die Informationen zum Team und den Öffnungszeiten finden könnten. Weitere vier Personen zeigten Verwirrung beim Anwenden der Filter, insbesondere beim Auffinden spezieller Produkte wie dem Matcha-Set oder der Etagere. Eine Person war enttäuscht, als die Suchfunktion nicht funktionierte.

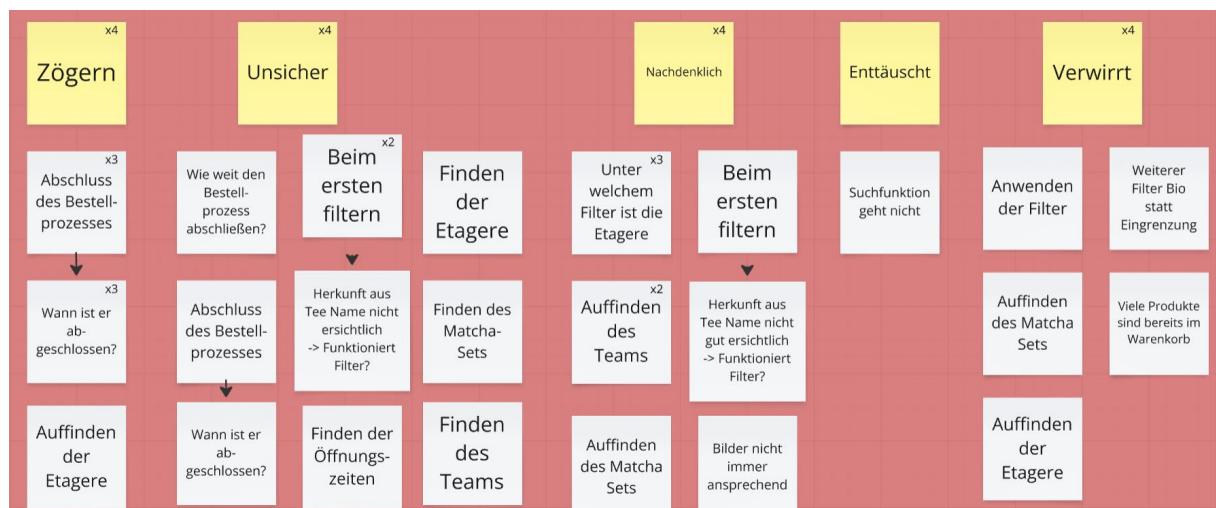


Abbildung 129. Verhalten während des Nutzens der manuell gestalteten Applikation

#### 4.4 Abschluss der Aufgaben

Insgesamt nahmen jeweils einundzwanzig Testpersonen pro Applikation an den Usability-Tests teil. Bei der KI-generierten Applikation konnten zwei Teilnehmer nicht alle Aufgaben erfolgreich abschließen. Die erste Person scheiterte an zwei der insgesamt acht Aufgaben. Grund hierfür war, dass sie den chinesischen Schwarztée sowie den aromatisierten Weißtee nicht finden konnte, da sie die Banner nicht gesehen hat, die Tees ohne weitere Eingrenzung jedoch nicht finden konnte und die Aufgabe eigenständig abbrach. Die zweite Testperson konnte die Aufgabe, einen Bio-Kräutertee zu finden, nicht erfüllen. Sie bemerkte nicht, dass der Filter Bio alle Bio-Tees der verschiedenen Sorten angezeigt. Statt dessen ging sie davon aus, dass es sich bei den Tees ausschließlich um Kräutertees handele und wählte daraufhin einen falschen Tee aus. Aufgrund dieser Fehler wurde die Aufgabe als nicht erfüllt angesehen.

Auch bei der manuell gestalteten Applikation konnten zwei von einundzwanzig Teilnehmern nicht alle Aufgaben erfolgreich abschließen. Im ersten Fall glaubte die Testperson, die Aufgabe des Abschlusses des Bestellprozesses nicht erfüllt zu haben, da sie nicht erkannte, dass die Bestellung bereits abgeschlossen war. Im Grunde hat sie die Aufgabe zwar abgeschlossen, da die Testperson dies jedoch anders wahrnahm, wurde die Aufgabe dennoch als nicht erfüllt bewertet. Die zweite Person wendete bei der Aufgabe, einen chinesischen Schwarztée zu finden, nach der Auswahl des Filters Schwarztée

keinen weiteren Filter an. Dadurch wurde anstelle eines chinesischen Tees ein indischer Schwarztee ausgewählt, weshalb die Aufgabe ebenfalls als nicht erfolgreich abgeschlossen gewertet wurde.

## 5. Interpretation der Ergebnisse

In den folgenden Unterkapiteln werden die Ergebnisse des Mann-Whitney-U-Tests sowie der deskriptiven Statistik betrachtet und interpretiert. Die Datensätze wurden dabei aufgeteilt in statistisch nicht signifikant und statistisch signifikant.

### 5.1 Statistisch nicht signifikante Datensätze

#### Schnell (1) / Langsam (7)

Da der Wert  $p = 0,0524$  größer als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch nicht signifikant. Es besteht also kein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als schnell oder langsam. Ein zufällig entstandener Unterschied kann dabei nicht ausgeschlossen werden. Der Wert für  $r = 0,3$  kann nach Cohen als mittlerer Effekt interpretiert werden und könnte damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar sein.

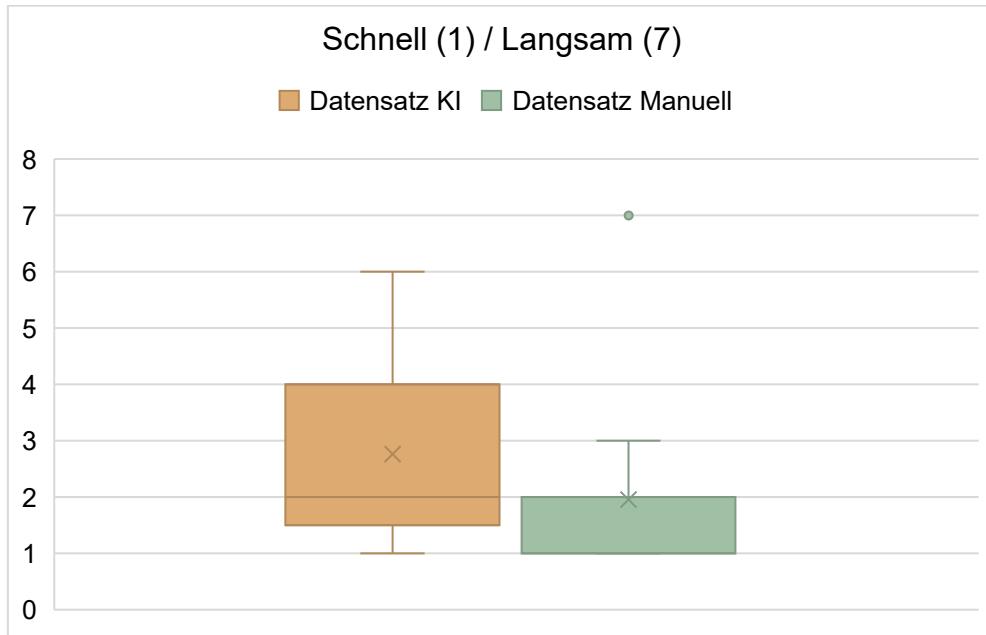


Abbildung 130. Boxplot-Diagramm für die Bewertung der Schnelligkeit

Die KI-Applikation hat einen Median von 2, was einer Bewertung im positiven Bereich entspricht. Der Interquartilsabstand von 2 zeigt eine leichte Streuung der Bewertungen an, womit die Einschätzung der Schnelligkeit unterschiedlich ausfiel. Auch der Mittelwert liegt mit 2,76 im positiven Bereich, weicht jedoch leicht vom Median ab und könnte auf Ausreißer in den Werten hinweisen.

Die manuelle Applikation weist mit einem Median von 2 und einem Interquartilsabstand von 2 darauf hin, dass die Applikation mit einer leichten Streuung als schnell bewertet wurde. Hier stimmen die

Bewertungen der Nutzer also nicht ganz überein. Auch der Mittelwert von 1,95 befindet sich im oberen positiven Bereich und bestätigt den Median.

Es wird jedoch deutlich, dass die beiden Applikationen sich bezüglich der Schnelligkeit nicht unterscheiden. Dies ist vor allem gut an dem Median zu sehen, der bei beiden Applikationen gleich ist. Dies bestätigt auch das Ergebnis des Mann-Whitney-U-Tests, der einen statistisch signifikanten Unterschied ablehnt. Das Boxplot-Diagramm könnte man zwar so interpretieren, dass es einen Unterschied gibt, da die zwei Visualisierungen der Werte durchaus unterschiedlich erscheinen. Dabei ist jedoch anzumerken, dass der Ausreißer bei der manuellen Applikation mit dem Wert von 7 den Mittelwert erhöht und damit verschlechtert. Weiterhin entstand durch den Mann-Whitney-U-Test nur knapp ein statistisch nicht signifikanter Unterschied. Das Ergebnis ist also empfindlich gegenüber Veränderungen.

### Leicht zu lernen (1) / Schwer zu lernen (7)

Da der Wert  $p = 0,0734$  größer als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch nicht signifikant. Es besteht also kein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als leicht zu lernen oder schwer zu lernen. Ein zufällig entstandener Unterschied kann dabei nicht ausgeschlossen werden. Der Wert für  $r = 0,28$  kann nach Cohen als kleiner Effekt interpretiert werden und bestätigt ebenfalls, dass der Unterschied zwischen den beiden Applikationen relativ gering ist.

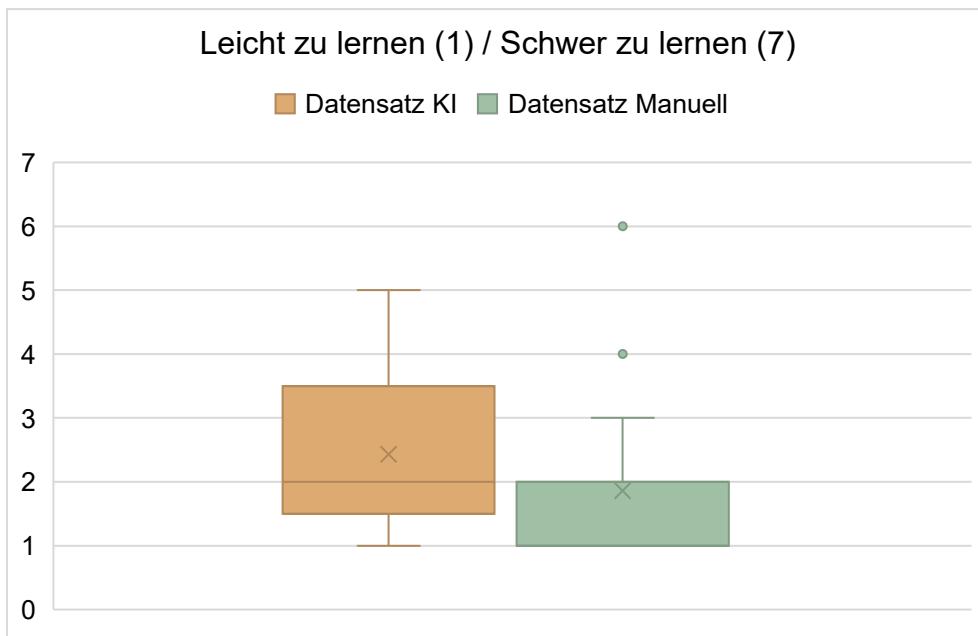


Abbildung 131. Boxplot-Diagramm für die Bewertung der Erlernbarkeit

Die KI-Applikation hat einen Median von 2, was einer Bewertung im positiven Bereich entspricht. Der Interquartilsabstand von 1 deutet auf eine einheitliche Einschätzung der Erlernbarkeit hin. Auch der Mittelwert liegt mit 2,43 im positiven Bereich.

Die manuelle Applikation weist mit einem Median von 1 und einem Interquartilsabstand von 1 darauf hin, dass die Applikation mit einheitlich als schnell bewertet wurde. Auch der Mittelwert von 1,86 befindet

sich im oberen positiven Bereich, unterscheidet sich jedoch leicht vom Median und könnte ein Hinweis auf Ausreißer sein.

Es wird jedoch deutlich, dass die beiden Applikationen sich bezüglich der Erlernbarkeit kaum unterscheiden. Dies ist gut an dem Mittelwert zu sehen, der bei beiden Applikationen nah beieinander liegt. Dies bestätigt auch das Ergebnis des Mann-Whitney-U-Tests, der einen statistisch signifikanten Unterschied ablehnt. Das Boxplot-Diagramm könnte man zwar so interpretieren, dass es einen Unterschied gibt, da die zwei Visualisierungen der Werte durchaus unterschiedlich erscheinen. Dabei ist jedoch anzumerken, dass die zwei Ausreißer bei der manuellen Applikation mit den Werten 4 und 6 den Mittelwert erhöhen und damit verschlechtern. Jedoch entstand durch den Mann-Whitney-U-Test ein deutlicher statistischer nicht signifikanter Unterschied. Das Ergebnis ist also dennoch nicht statistisch signifikant.

### Aktivierend (1) / Einschläfernd (7)

Da der Wert  $p = 0,0802$  größer als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch nicht signifikant. Es besteht also kein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als aktivierend oder einschläfernd. Ein zufällig entstandener Unterschied kann dabei nicht ausgeschlossen werden. Der Wert für  $r = 0,27$  kann nach Cohen als kleiner Effekt interpretiert werden und bestätigt ebenfalls, dass der Unterschied zwischen den beiden Applikationen relativ gering ist.

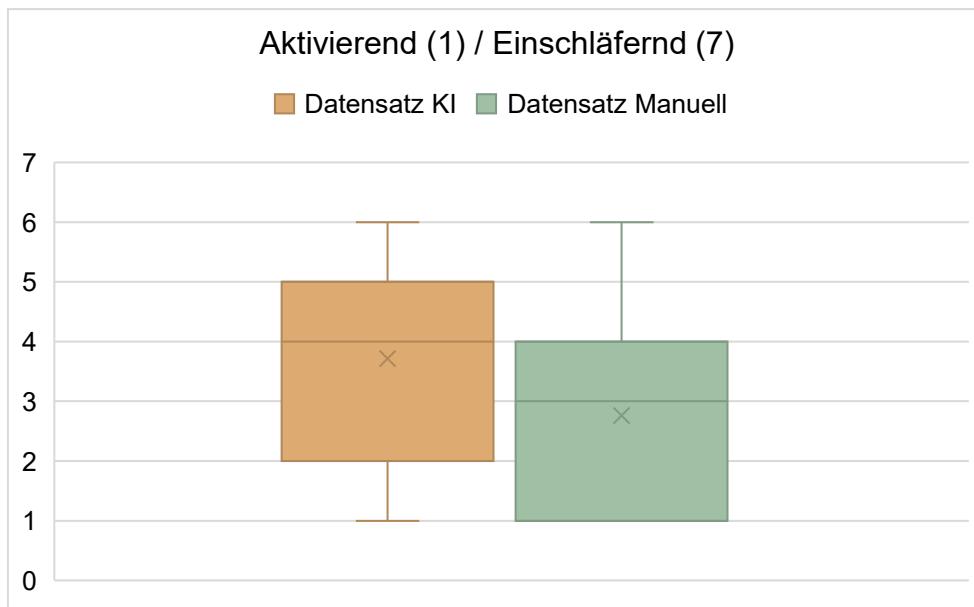


Abbildung 132. Boxplot-Diagramm für die Bewertung der Stimulierung

Am Boxplot-Diagramm zeigt sich, dass der Unterschied bei der Bewertung der Stimulierung zwischen der KI-generierten Applikation und der manuell erstellten Applikation gering ist.

Die KI-Applikation hat einen Median von 4, was einer Bewertung im neutralen Bereich entspricht. Der starke Interquartilsabstand von 3 zeigt eine starke Streuung der Bewertungen an, womit die Einschätzung sehr unterschiedlich ausfiel. Auch der Mittelwert liegt mit 3,71 annähernd im neutralen Bereich.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 2 und einem Interquartilsabstand von 3 darauf hin, dass die Applikation mit einer hohen Streuung als aktivierend bewertet wurde. Hier stimmen die Bewertungen der Nutzer also nicht ganz überein. Der Mittelwert von 2,76 weicht vom Median ab und könnte ein Hinweis auf Ausreißer sein.

Trotz des unterschiedlichen Medians beider Applikationen ist der Unterschied statistisch nicht signifikant, da die Bewertungen stark gestreut sind und sich zu oft überlappen.

### Kreativ (1) / Phantasielos (7)

Da der Wert  $p = 0,097$  größer als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch nicht signifikant. Es besteht also kein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als kreativ oder phantasielos. Ein zufällig entstandener Unterschied kann dabei nicht ausgeschlossen werden. Der Wert für  $r = 0,26$  kann nach Cohen als kleiner Effekt interpretiert werden und bestätigt ebenfalls, dass der Unterschied zwischen den beiden Applikationen relativ gering ist.

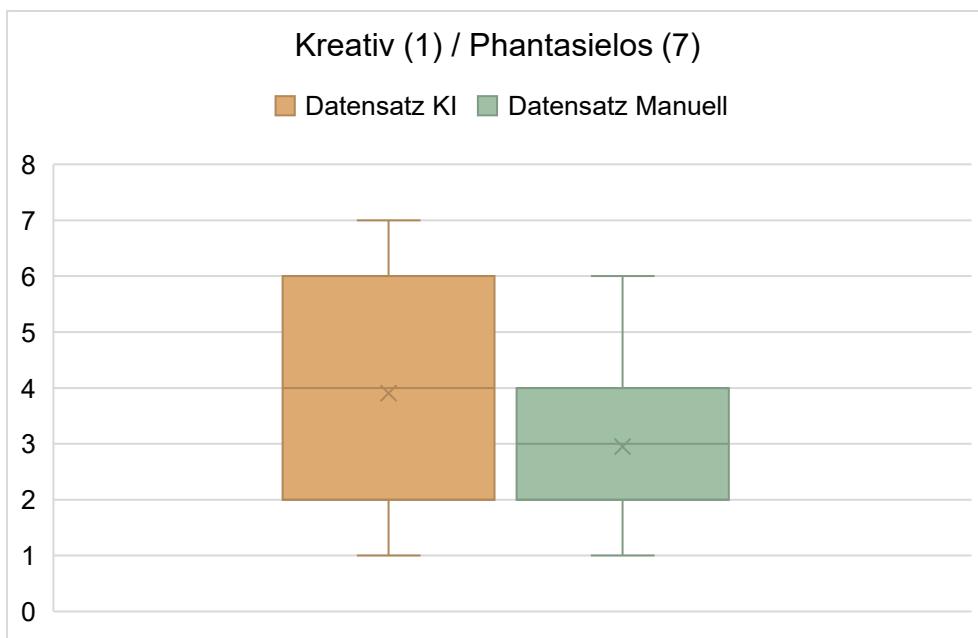


Abbildung 133. Boxplot-Diagramm für die Bewertung der Kreativität

Am Boxplot-Diagramm zeigt sich, dass der Unterschied bei der Bewertung der Kreativität zwischen der KI-generierten Applikation und der manuell erstellten Applikation gering ist.

Die KI-Applikation hat einen Median von 4, was einer Bewertung im neutralen Bereich entspricht. Der Interquartilsabstand von 4 zeigt eine sehr starke Streuung der Bewertungen an, womit die Einschätzung sehr unterschiedlich ausfiel. Auch der Mittelwert liegt mit 3,9 annähernd im neutralen Bereich.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 3 und einem Interquartilsabstand von 2 darauf hin, dass die Applikation mit einer leichten Streuung als kreativ bewertet wurde. Hier stimmen die Bewertungen der Nutzer also nicht ganz überein. Der Mittelwert von 2,95 bestätigt den Median.

Trotz des unterschiedlichen Medians beider Applikationen ist der Unterschied statistisch nicht signifikant, da die Bewertungen stark gestreut sind und sich zu oft überlappen.

### Sicher (1) / Unsicher (7)

Da der Wert  $p = 0,2006$  größer als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch nicht signifikant. Es besteht also kein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als sicher oder unsicher. Ein zufällig entstandener Unterschied kann dabei nicht ausgeschlossen werden. Der Wert für  $r = 0,2$  kann nach Cohen als kleiner Effekt interpretiert werden und bestätigt ebenfalls, dass der Unterschied zwischen den beiden Applikationen relativ gering ist.

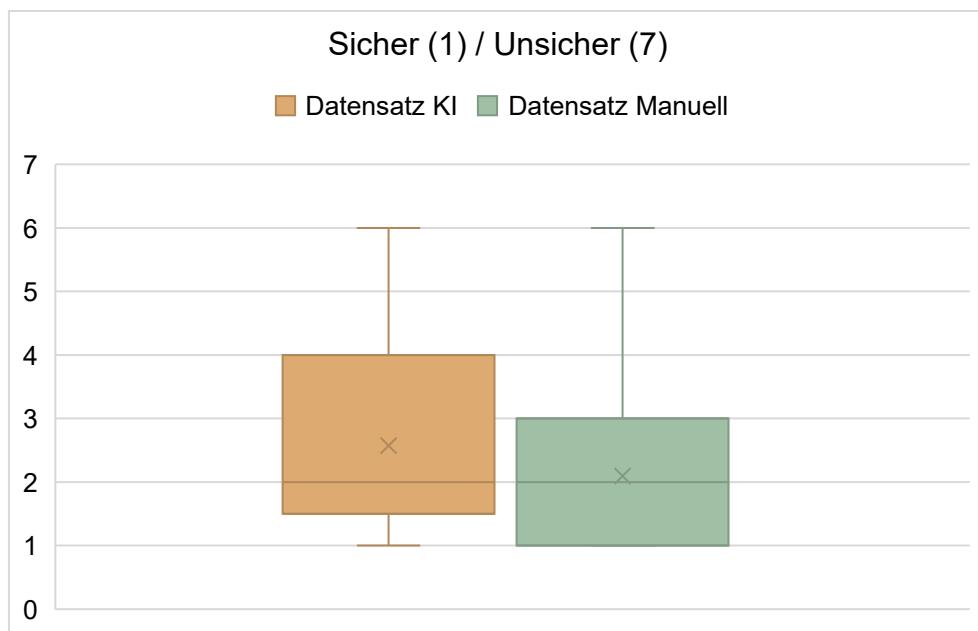


Abbildung 134. Boxplot-Diagramm für die Bewertung der Sicherheit

Am Boxplot-Diagramm zeigt sich, dass der Unterschied bei der Bewertung der Sicherheit zwischen der KI-generierten Applikation und der manuell erstellten Applikation gering ist.

Die KI-Applikation hat einen Median von 2, was einer Bewertung im positiven Bereich entspricht. Der Interquartilsabstand von 2 zeigt eine leichte Streuung der Bewertungen an, womit die Einschätzung unterschiedlich ausfiel. Auch der Mittelwert liegt mit 2,57 im positiven Bereich, weicht jedoch leicht vom Median ab und könnte auf Ausreißer hinweisen.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 2 und einem Interquartilsabstand von 2 darauf hin, dass die Applikation mit einer leichten Streuung als sicher bewertet wurde. Hier stimmen die Bewertungen der Nutzer also nicht ganz überein. Der Mittelwert von 2,1 bestätigt den Median.

Der gleiche Median beider Applikationen sowie der ähnliche Mittelwert bei einer geringen Streuung beweist erneut, dass der Unterschied statistisch nicht signifikant ist.

## 5.2 Statistisch signifikante Datensätze

### Übersichtlich (1) / Verwirrend (7)

Da der Wert  $p = 0,0002$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als übersichtlich oder verwirrend. Der Unterschied ist dadurch nicht zufällig entstanden. Der Wert für  $r = 0,63$  kann nach Cohen als großer Effekt interpretiert werden und kennzeichnet einen deutlichen Unterschied zwischen den Applikationen, bei denen sich die Werte kaum überlappen.

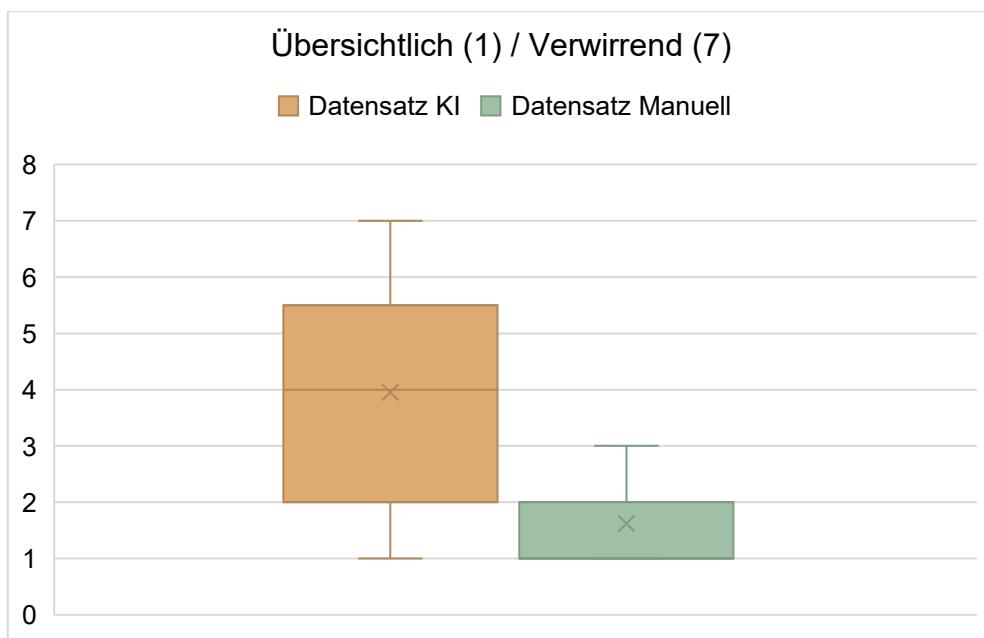


Abbildung 135. Boxplot-Diagramm für die Bewertung der Übersichtlichkeit

Am Boxplot-Diagramm zeigt sich ein deutlicher Unterschied bei der Bewertung der Übersichtlichkeit zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 4, was einer Bewertung im neutralen Bereich entspricht. Der Interquartilsabstand von 3 zeigt eine starke Streuung der Bewertungen an, womit die Einschätzung der Übersichtlichkeit sehr unterschiedlich ausfiel. Auch der Mittelwert liegt mit 3,95 annähernd im neutralen Bereich.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 1 und einem Interquartilsabstand von 1 darauf hin, dass die Applikation einheitlich als sehr übersichtlich bewertet wurde. Hier stimmen die Bewertungen der Nutzer also weitgehend überein. Auch der Mittelwert von 1,62 befindet sich im oberen positiven Bereich und bestätigt diese Beobachtung.

### **Behindernd (1) / Unterstützend (7)**

Da der Wert  $p = 0,0002$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als behindernd oder unterstützend. Der Unterschied ist nicht zufällig entstanden. Der Wert für  $r = 0,63$  kann nach Cohen als großer Effekt interpretiert werden und kennzeichnet einen deutlichen Unterschied zwischen den Applikationen, bei denen sich die Werte kaum überlappen.

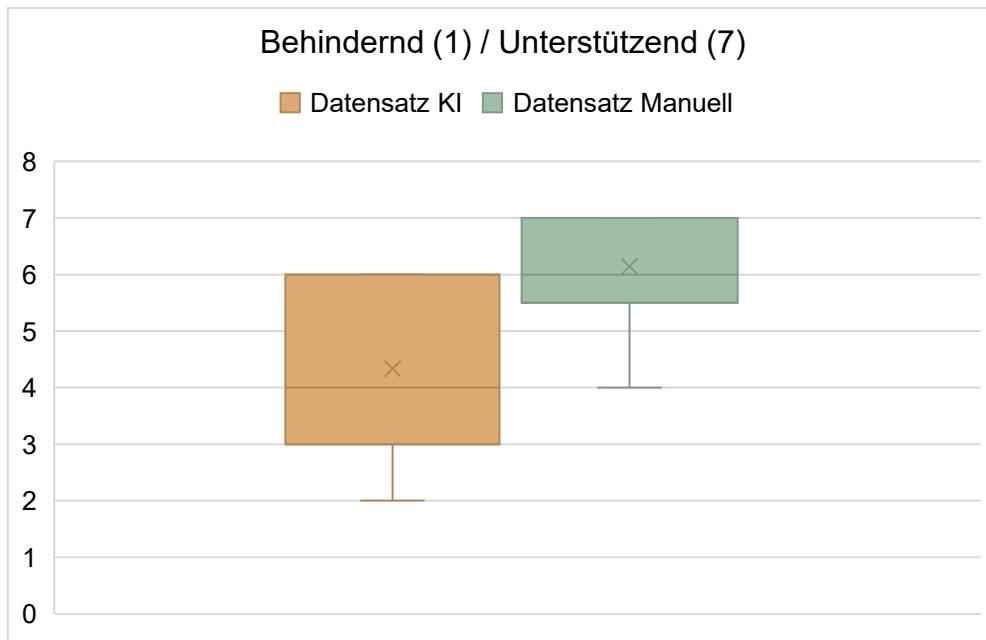


Abbildung 136. Boxplot-Diagramm für die Bewertung der Unterstützung

Am Boxplot-Diagramm zeigt sich ein deutlicher Unterschied bei der Bewertung der Unterstützung zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 4, was einer Bewertung im neutralen Bereich entspricht. Der Interquartilsabstand von 3 zeigt eine starke Streuung der Bewertungen an, womit die Einschätzung der Unterstützung sehr unterschiedlich ausfiel. Auch der Mittelwert liegt mit 4,33 im mittleren neutralen Bereich der Skala.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 6 und einem Interquartilsabstand von 1 darauf hin, dass die Applikation einheitlich als sehr unterstützend bewertet wurde. Hier stimmen die Bewertungen der Nutzer also weitgehend überein. Auch der Mittelwert von 6,14 befindet sich im oberen positiven Bereich und bestätigt diese Beobachtung.

### **Unverständlich (1) / Verständlich (7)**

Da der Wert  $p = 0,0002$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als unverständlich oder verständlich. Der Unterschied ist nicht zufällig entstanden. Der Wert für  $r = 0,62$  kann nach Cohen als großer Effekt interpretiert werden und kennzeichnet einen deutlichen Unterschied zwischen den Applikationen, bei denen sich die Werte kaum überlappen.

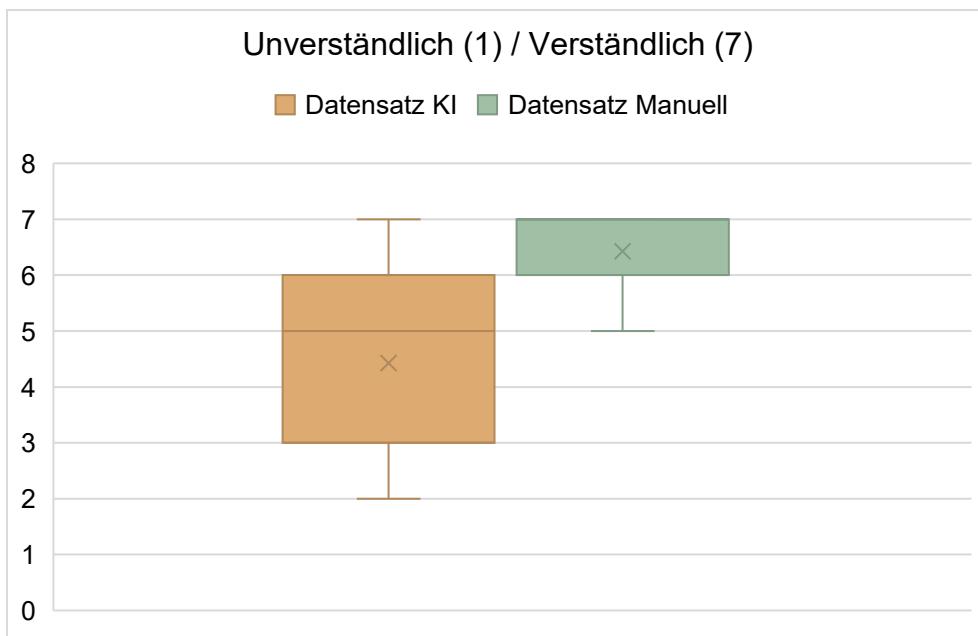


Abbildung 137. Boxplot-Diagramm für die Bewertung der Verständlichkeit

Am Boxplot-Diagramm zeigt sich ein deutlicher Unterschied bei der Bewertung der Verständlichkeit zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 5, was einer Bewertung im positiven Bereich entspricht. Der Interquartilsabstand von 3 zeigt eine starke Streuung der Bewertungen an, womit die Einschätzung der Verständlichkeit sehr unterschiedlich ausfiel. Der Mittelwert liegt mit 4,43 eher im mittleren neutralen Bereich der Skala. Die leichte Abweichung des Medians im Vergleich zu dem Mittelwert könnte auf Ausreißer in den Werten hindeuten.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 7 und einem Interquartilsabstand von 1 darauf hin, dass die Applikation einheitlich als sehr verständlich bewertet wurde. Hier stimmen die Bewertungen der Nutzer also weitgehend überein. Auch der Mittelwert von 6,43 befindet sich im oberen positiven Bereich und bestätigt diese Beobachtung. Die leichte Abweichung des Medians im Vergleich zu dem Mittelwert könnte auch hier auf Ausreißer in den Werten hinweisen.

### Ineffizient (1) / Effizient (7)

Da der Wert  $p = 0,0006$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als ineffizient oder effizient. Der Unterschied ist dadurch nicht zufällig entstanden. Der Wert für  $r = 0,53$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

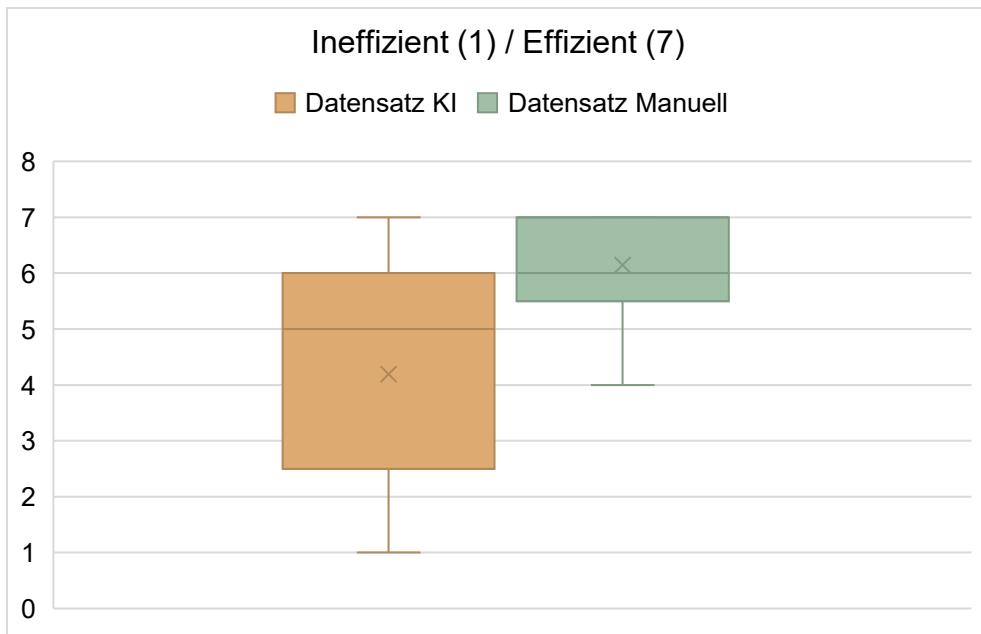


Abbildung 138. Boxplot-Diagramm für die Bewertung der Effizienz

Am Boxplot-Diagramm zeigt sich ein deutlicher Unterschied bei der Bewertung der Effizienz zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 5, was einer Bewertung im positiven Bereich entspricht. Der Interquartilsabstand von 3 zeigt eine starke Streuung der Bewertungen an, womit die Einschätzung der Effizienz sehr unterschiedlich ausfielen. Der Mittelwert liegt mit 4,19 eher im mittleren neutralen Bereich der Skala. Die Abweichung des Medians im Vergleich zu dem Mittelwert könnte auf Ausreißer in den Werten hindeuten.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 6 und einem Interquartilsabstand von 1 darauf hin, dass die Applikation einheitlich als sehr effizient bewertet wurde. Hier stimmen die Bewertungen der Nutzer also weitgehend überein. Auch der Mittelwert von 6,14 befindet sich im oberen positiven Bereich und bestätigt diese Beobachtung.

### **Sehr Gut (1) / Schlecht (7)**

Da der Wert  $p = 0,0012$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als sehr gut oder schlecht. Der Unterschied ist nicht zufällig entstanden. Der Wert für  $r = 0,50$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

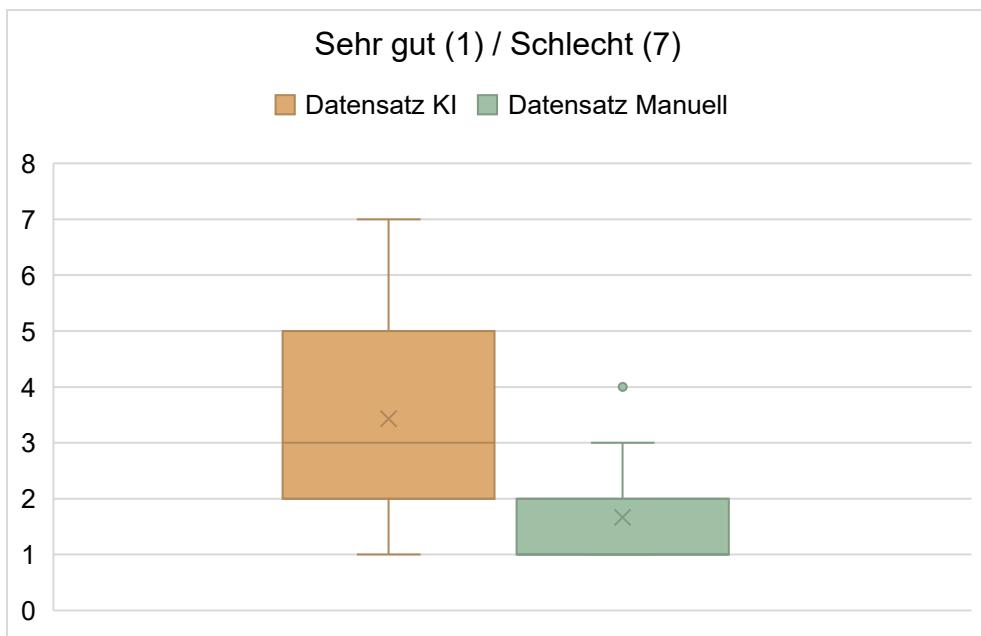


Abbildung 139. Boxplot-Diagramm für die Bewertung des Gefallens

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung des Gefallens zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 3, was einer Bewertung im positiven Bereich entspricht. Der Interquartilsabstand von 3 zeigt eine starke Streuung der Bewertungen an, womit die Einschätzung des Gefallens sehr unterschiedlich ausfiel. Der Mittelwert liegt mit 3,43 eher im neutralen bis positiven Bereich der Skala.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 1 und einem Interquartilsabstand von 1 darauf hin, dass die Applikation einheitlich als sehr gut bewertet wurde. Hier stimmen die Bewertungen der Nutzer also weitgehend überein. Allerdings gibt es dennoch einen Ausreißer nach oben mit einer Bewertung von 4. Auch der Mittelwert von 1,66 befindet sich im oberen positiven Bereich und bestätigt diese Beobachtung.

## Unerfreulich (1) / Erfreulich (7)

Da der Wert  $p = 0,0018$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als unerfreulich oder erfreulich. Der Unterschied ist dadurch nicht zufällig entstanden. Der Wert für  $r = 0,48$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

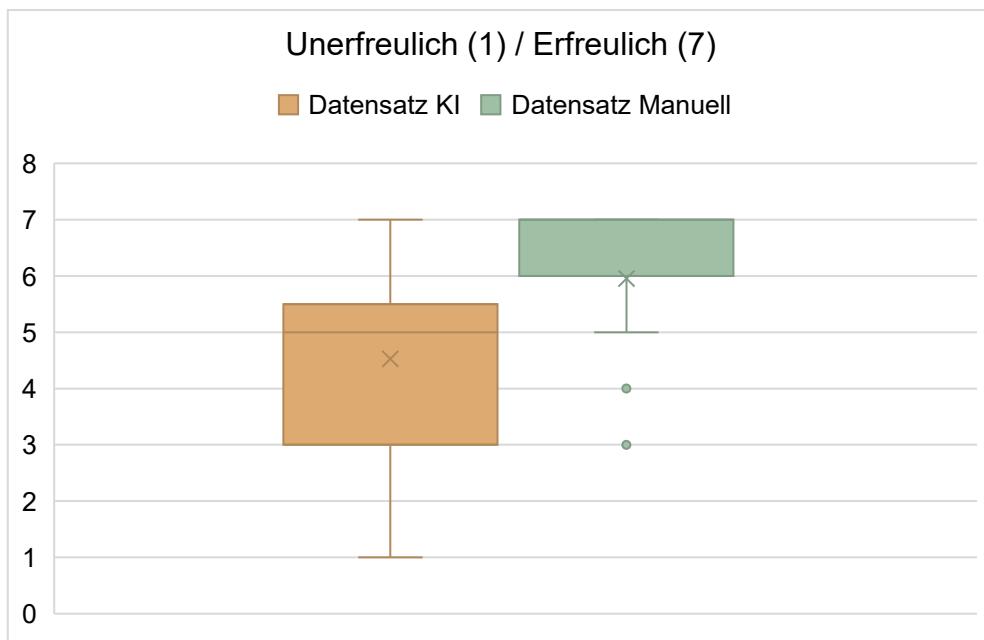


Abbildung 140. Boxplot-Diagramm für die Bewertung der Erfreulichkeit

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung der Erfreulichkeit zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 5, was einer Bewertung im positiven Bereich entspricht. Der Interquartilsabstand von 2 zeigt eine leichte Streuung der Bewertungen an, womit die Einschätzung der Erfreulichkeit eher unterschiedlich ausfiel. Der Mittelwert liegt mit 4,52 eher im positiven bis neutralen Bereich der Skala und weicht leicht vom Median ab. Dies könnte ein Hinweis auf Ausreißer sein.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 6 und einem Interquartilsabstand von 1 darauf hin, dass die Applikation einheitlich als sehr erfreulich angesehen wurde. Hier stimmen die Bewertungen der Nutzer also weitgehend überein. Allerdings gibt es dennoch Ausreißer nach unten mit einer Bewertung von 3 und 4. Auch der Mittelwert von 5,95 befindet sich im oberen positiven Bereich und bestätigt diese Beobachtung.

## Abstoßend (1) / Anziehend (7)

Da der Wert  $p = 0,0018$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als abstoßend oder anziehend. Der Unterschied ist dadurch nicht zufällig entstanden. Der Wert für  $r = 0,48$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

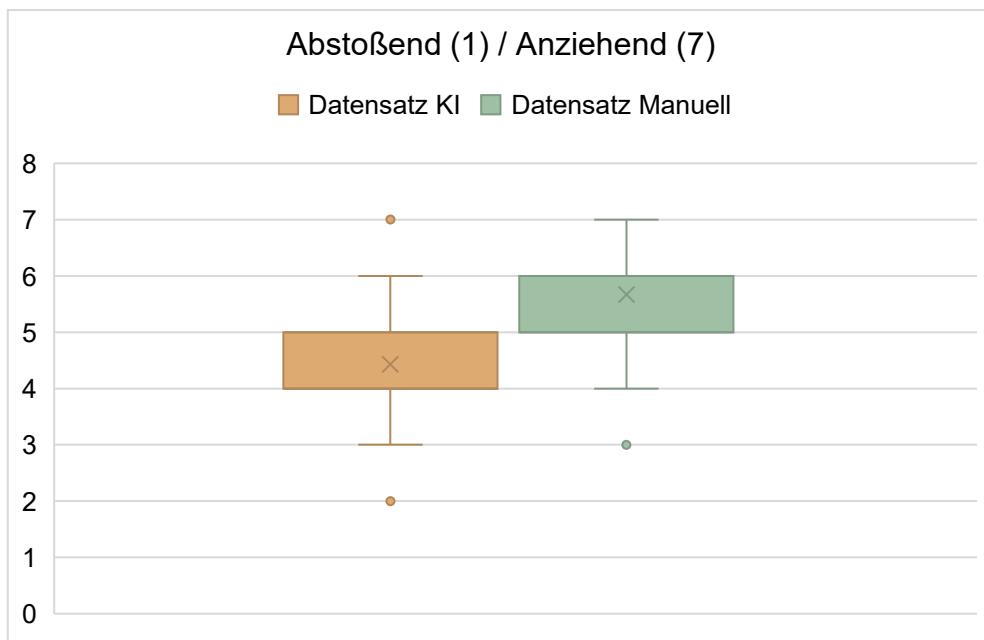


Abbildung 141. Boxplot-Diagramm für die Bewertung der Anziehung

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung der Anziehung zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 4, was einer Bewertung im neutralen Bereich entspricht. Der Interquartilsabstand von 1 zeigt eine geringe Streuung der Bewertungen an, womit die Einschätzung der Anziehung sehr einheitlich ausfiel. Allerdings gibt es dennoch Ausreißer sowohl nach unten mit einer Bewertung von 2 als auch nach oben mit einer Bewertung von 7. Der Mittelwert liegt mit 4,43 eher im neutralen Bereich der Skala und bestätigt den Median.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 6 und einem Interquartilsabstand von 1 darauf hin, dass die Applikation einheitlich als sehr anziehend bewertet wurde. Hier stimmen die Bewertungen der Nutzer also weitgehend überein. Jedoch gibt es einen Ausreißer im unteren Bereich mit einer Bewertung von 3. Auch der Mittelwert von 5,67 befindet sich im oberen positiven Bereich und bestätigt die Beobachtung des Medians.

### **Uninteressant (1) / Interessant (7)**

Da der Wert  $p = 0,002$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als uninteressant oder interessant. Der Unterschied ist nicht zufällig entstanden. Der Wert für  $r = 0,48$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

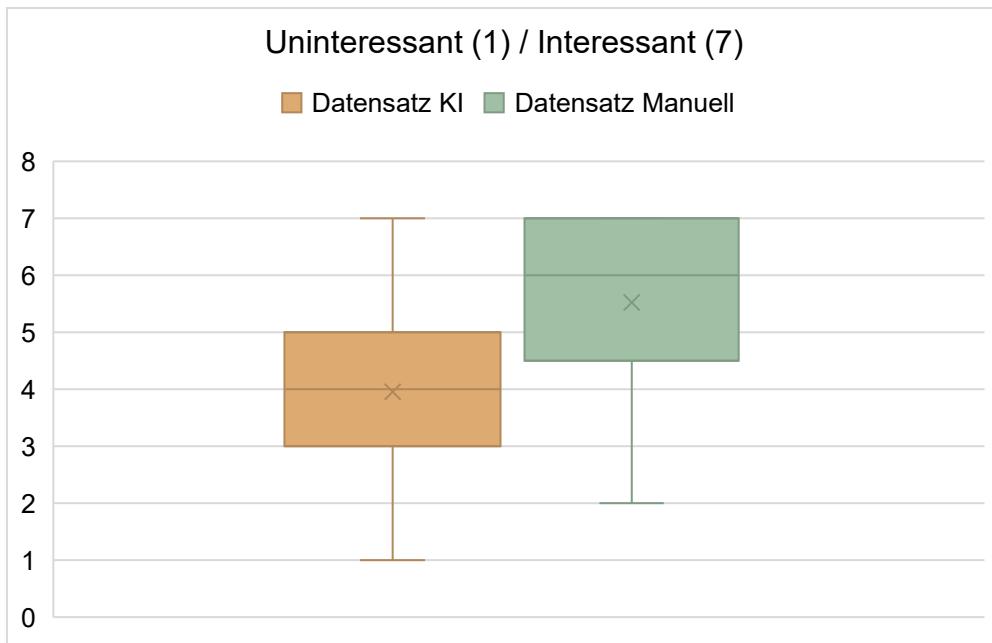


Abbildung 142. Boxplot-Diagramm für die Bewertung des Interesses

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung des Interesses zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 4, was einer Bewertung im neutralen Bereich entspricht. Der Interquartilsabstand von 2 zeigt eine leicht erhöhte Streuung der Bewertungen an, womit die Einschätzung des Interesses nicht ganz einheitlich ausfiel. Der Mittelwert liegt mit 3,95 annähernd im neutralen Bereich und bestätigt den Median.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 6 und einem Interquartilsabstand von 2 darauf hin, dass die Applikation nicht ganz einheitlich als interessant bewertet wurde. Hier stimmen die Bewertungen der Nutzer also nur leicht überein. Auch der Mittelwert von 5,52 befindet sich im oberen positiven Bereich und bestätigt die Beobachtung des Medians.

### **Navigation: sehr einfach (1) / sehr schwierig (7)**

Da der Wert  $p = 0,002$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung der Navigation als sehr einfach oder sehr schwierig. Der Unterschied ist nicht zufällig entstanden. Der Wert für  $r = 0,48$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

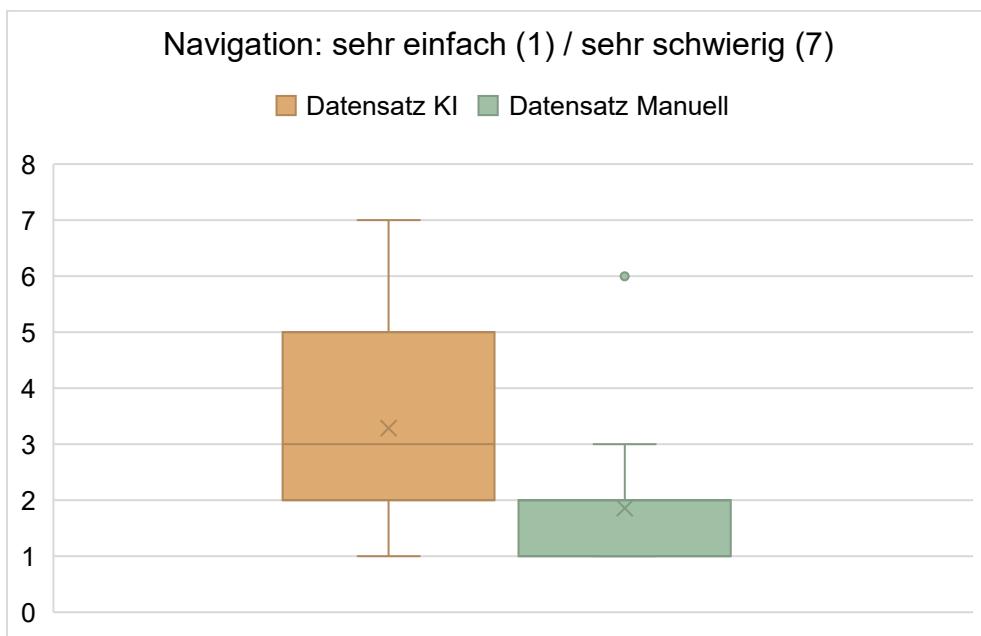


Abbildung 143. Boxplot-Diagramm für die Bewertung der Navigation

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung der Navigation zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 3, was einer positiven Bewertung entspricht. Der Interquartilsabstand von 3 zeigt eine starke Streuung der Bewertungen an, womit die Einschätzung der Navigation nicht einheitlich ausfiel. Der Mittelwert liegt mit 3,29 im positiven Bereich und bestätigt den Median.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 2 und einem Interquartilsabstand von 1 darauf hin, dass die Navigation der Applikation einheitlich als gut bewertet wurde. Hier stimmen die Bewertungen der Nutzer also überein. Jedoch gibt es einen Ausreißer im oberen Bereich mit einer Bewertung von 6. Auch der Mittelwert von 1,86 befindet sich im oberen positiven Bereich und bestätigt die Beobachtung des Medians.

## Unangenehm (1) / Angenehm (7)

Da der Wert  $p = 0,0046$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als unangenehm oder angenehm. Der Unterschied ist dadurch nicht zufällig entstanden. Der Wert für  $r = 0,44$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

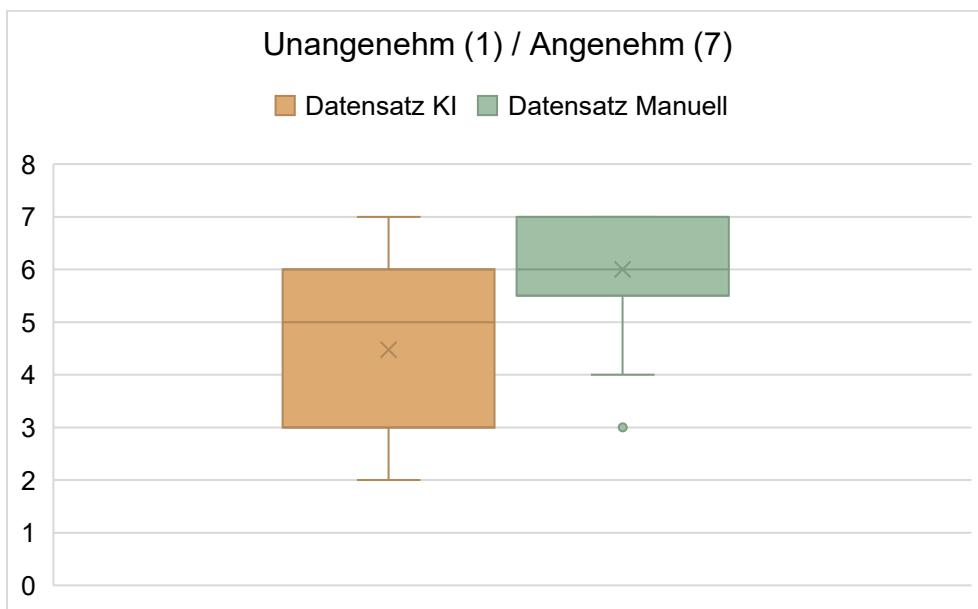


Abbildung 144. Boxplot-Diagramm für die Bewertung des Wohlbefindens

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung des Wohlbefindens zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 5, was einer Bewertung im positiven Bereich entspricht. Der Interquartilsabstand von 3 zeigt eine erhöhte Streuung der Bewertungen an, womit die Einschätzung des Wohlbefindens nicht ganz einheitlich ausfiel. Der Mittelwert liegt mit 4,48 annähernd im neutralen Bereich und weicht damit leicht von dem Median ab. Dies könnte auf Ausreißer in den Werten hindeuten.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 6 und einem Interquartilsabstand von 1 darauf hin, dass die Applikation einheitlich als sehr angenehm bewertet wurde. Hier stimmen die Bewertungen der Nutzer also überein. Jedoch gibt es einen Ausreißer im unteren Bereich mit einer Bewertung von 3. Auch der Mittelwert von 6 befindet sich im oberen positiven Bereich und bestätigt die Beobachtung des Medians.

### **Langweilig (1) / Spannend (7)**

Da der Wert  $p = 0,005$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als langweilig oder spannend. Der Unterschied ist dadurch nicht zufällig entstanden. Der Wert für  $r = 0,43$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

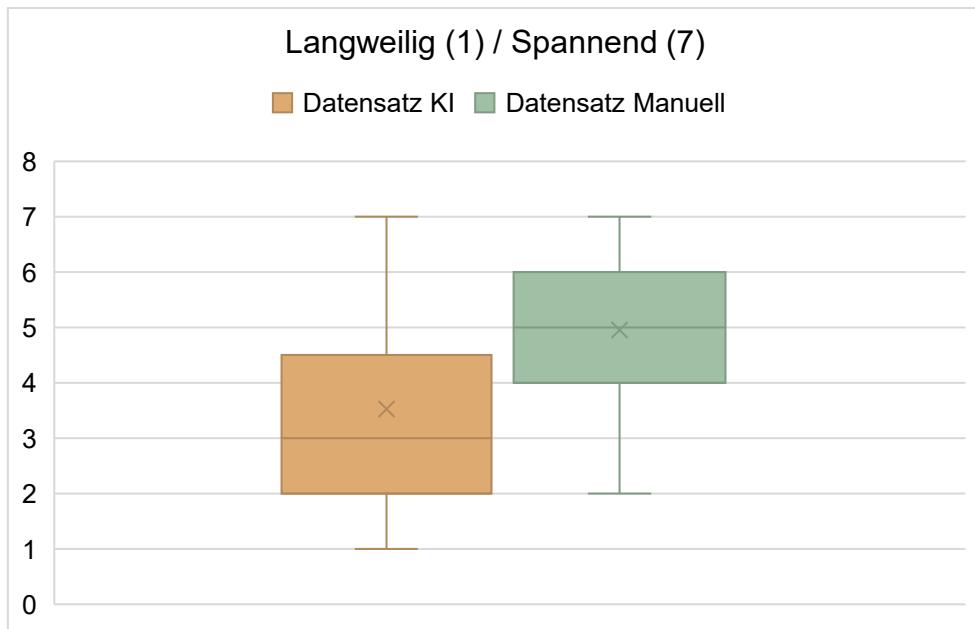


Abbildung 145. Boxplot-Diagramm für die Bewertung der Spannung

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung der Spannung zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 3, was einer Bewertung im negativen Bereich entspricht. Der Interquartilsabstand von 2 zeigt eine leicht erhöhte Streuung der Bewertungen an, womit die Einschätzung der Spannung nicht ganz einheitlich ausfiel. Der Mittelwert liegt mit 3,52 annähernd im neutralen Bereich und weicht leicht vom Median ab. Dies könnte auf leichte Ausreißer in der Bewertung hinweisen.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 5 und einem Interquartilsabstand von 2 darauf hin, dass die Applikation nicht ganz einheitlich als spannend bewertet wurde. Hier stimmen die Bewertungen der Nutzer also nicht ganz überein. Auch der Mittelwert von 4,95 befindet sich im oberen positiven Bereich und bestätigt die Beobachtung des Medians.

## **Erwartungskonform (1) / Nicht erwartungskonform (7)**

Da der Wert  $p = 0,0066$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als erwartungskonform oder nicht erwartungskonform. Der Unterschied ist nicht zufällig entstanden. Der Wert für  $r = 0,42$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

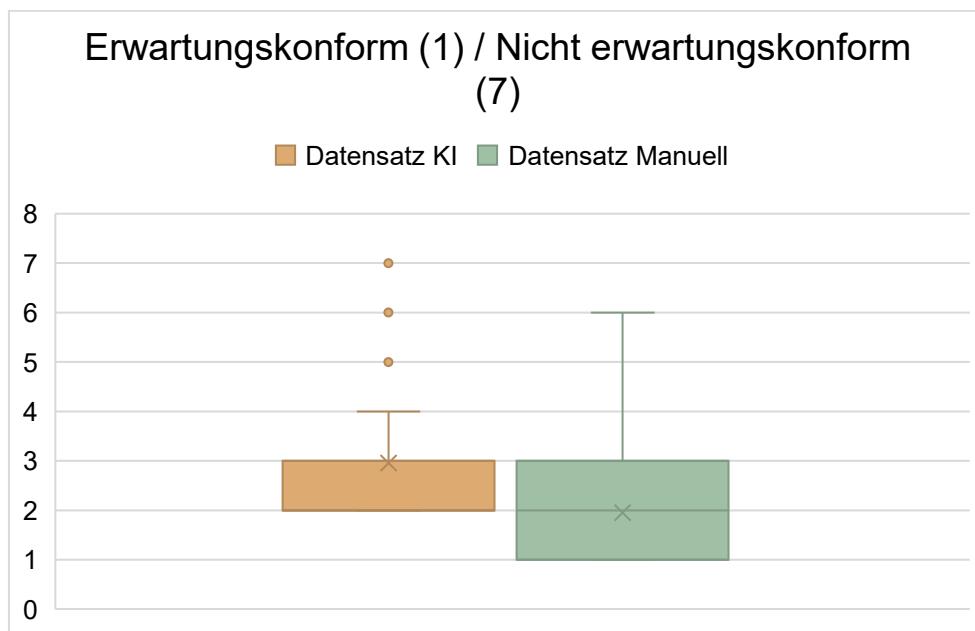


Abbildung 146. Boxplot-Diagramm für die Bewertung der Erwartungskonformität

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung der Erwartungskonformität zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 4, was einer Bewertung im neutralen Bereich entspricht. Der Interquartilsabstand von 3 zeigt eine erhöhte Streuung der Bewertungen an, womit die Einschätzung der Erwartungskonformität nicht ganz einheitlich ausfiel. Der Mittelwert liegt mit 2,95 im positiven Bereich und weicht damit deutlich von dem Median ab. Dies deutet auf Ausreißer in den Werten hin, die auch in dem Diagramm deutlich zu erkennen sind. Die Ausreißer nach oben liegen bei den Werten 5,6 und 7.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 3 und einem Interquartilsabstand von 3 darauf hin, dass die Applikation nicht einheitlich als erwartungskonform bewertet wurde. Hier stimmen die Bewertungen der Nutzer also nicht ganz überein. Der Mittelwert von 1,95 befindet sich im oberen positiven Bereich und weicht vom Median ab. Dies könnte auf Ausreißer hinweisen.

### Sortierung: Sehr gut (1) / Sehr schlecht (7)

Da der Wert  $p = 0,01$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung der Sortierung als sehr gut oder sehr schlecht. Der Unterschied ist nicht zufällig entstanden. Der Wert für  $r = 0,40$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

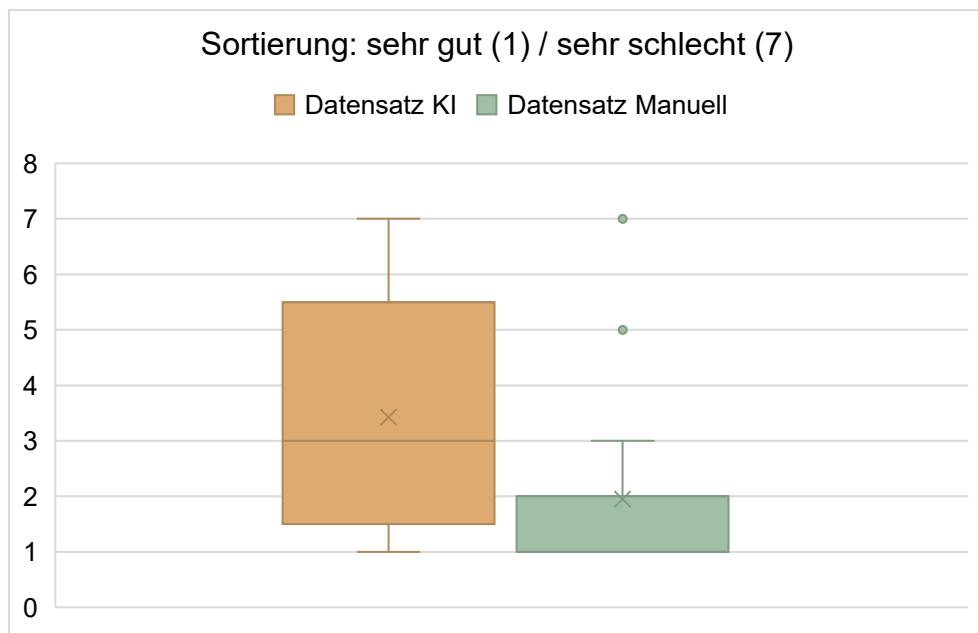


Abbildung 147. Boxplot-Diagramm für die Bewertung der Sortierung

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung der Sortierung zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 3, was einer Bewertung im positiven Bereich entspricht. Der Interquartilsabstand von 3 zeigt eine erhöhte Streuung der Bewertungen an, womit die Einschätzung der Sortierung nicht einheitlich ausfiel. Der Mittelwert liegt mit 3,43 im positiven bis neutralen Bereich.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 1 und einem Interquartilsabstand von 1 darauf hin, dass die Applikation einheitlich als gut sortiert bewertet wurde. Hier stimmen die Bewertungen der Nutzer also überein. Auch der Mittelwert von 1,95 befindet sich im oberen positiven Bereich, weicht jedoch vom Median ab und deutet auf Ausreißer hin. Dies ist auch im Diagramm zu erkennen, da sich im oberen Bereich zwei Ausreißer mit einer Bewertung von 5 und 7 befinden.

## Kompliziert (1) / Einfach (7)

Da der Wert  $p = 0,012$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als kompliziert oder einfach. Der Unterschied ist nicht zufällig entstanden. Der Wert für  $r = 0,39$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

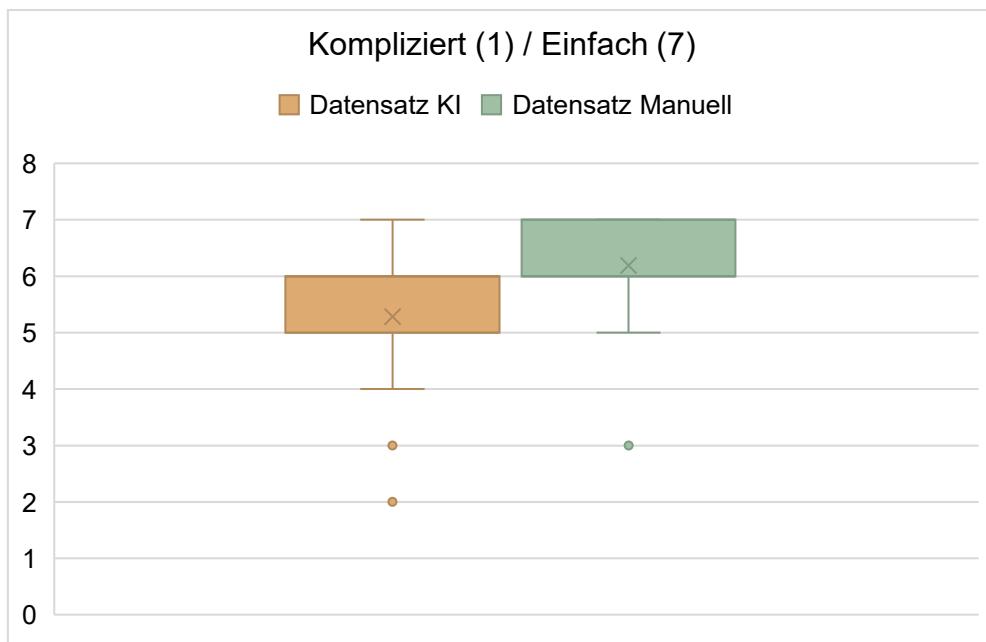


Abbildung 148. Boxplot-Diagramm für die Bewertung der Einfachheit

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung der Einfachheit zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 6, was einer Bewertung im positiven Bereich entspricht. Der Interquartilsabstand von 1 zeigt eine geringe Streuung der Bewertungen an, womit die Einschätzung der Einfachheit einheitlich ausfiel. Der Mittelwert liegt mit 5,29 im positiven Bereich, weicht vom Median ab und weist damit auf Ausreißer hin, die in dem Diagramm für die Werte 2 und 3 zu erkennen sind.

Die manuelle Applikation weist mit einem Median von 6 und einem Interquartilsabstand von 1 darauf hin, dass die Applikation einheitlich als einfach bewertet wurde. Hier stimmen die Bewertungen der Nutzer also überein. Auch der Mittelwert von 6,19 befindet sich im oberen positiven Bereich und bestätigt den Median.

Trotz des durch den Mann-Whitney-U-Test ermittelten statistisch signifikanten Unterschieds in der Einfachheit zwischen der KI-Applikation und der manuellen Applikation für  $p = 0,012$  und  $r = 0,39$  und dem erkennbaren Unterschied bei den Mittelwerten, ist beim Median kein Unterschied zu erkennen. Dennoch besteht ein Unterschied in der gesamten Verteilung der Werte, wie zum Beispiel in den Extremen oder durch Ausreißer. Diese sind auch in dem Diagramm zu erkennen. Bei der KI-generierten Applikation

befinden sich diese im unteren negativen Bereich mit den Werten 3 und 2. Auch die manuell gestaltete Applikation hat einen Ausreißer im unteren negativen Bereich mit dem Wert 3.

### Konservativ (1) / Innovativ (7)

Da der Wert  $p = 0,00124$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als konservativ oder innovativ. Der Unterschied ist damit nicht zufällig entstanden. Der Wert für  $r = 0,39$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

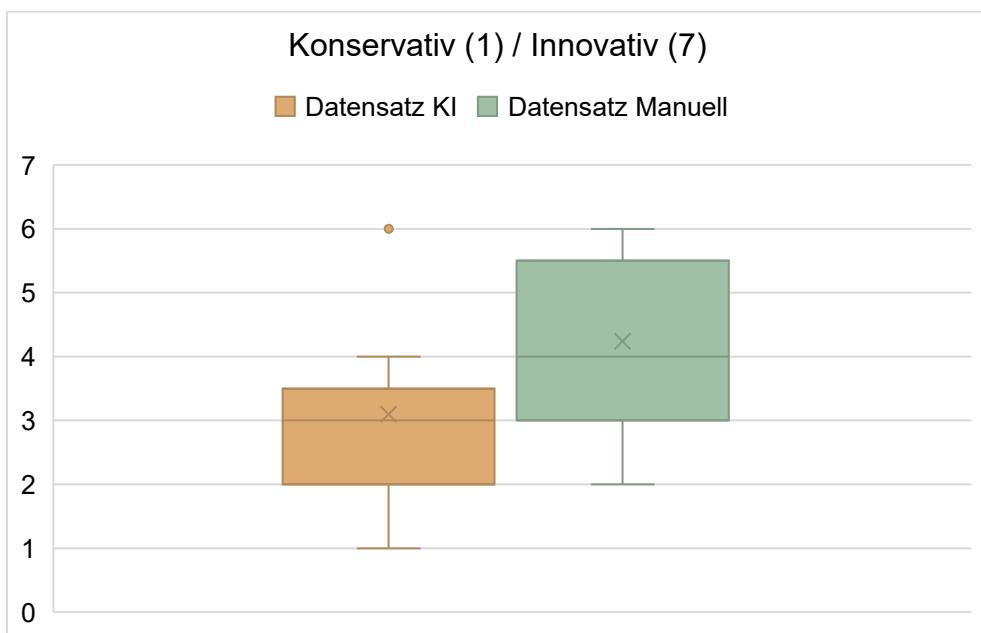


Abbildung 149. Boxplot-Diagramm für die Bewertung der Innovativität

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung der Innovativität zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 3, was einer Bewertung im negativen Bereich entspricht. Der Interquartilsabstand von 1 zeigt eine niedrige Streuung der Bewertungen an, womit die Einschätzung der Innovativität einheitlich ausfiel. Jedoch ist im Diagramm ein Ausreißer bei der Bewertung im oberen positiven Bereich mit dem Wert 6 zu erkennen. Der Mittelwert liegt mit 3,1 ebenfalls im negativen Bereich.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 4 und einem Interquartilsabstand von 2 darauf hin, dass die Applikation mit einer leichten Streuung als neutral bewertet wurde. Hier stimmen die Bewertungen der Nutzer also überein. Auch der Mittelwert von 4,24 befindet sich im mittleren neutralen Bereich und bestätigt den Median.

## Herkömmlich (1) / Neuartig (7)

Da der Wert  $p = 0,0136$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als herkömmlich oder neuartig. Der Unterschied ist damit nicht zufällig entstanden. Der Wert für  $r = 0,38$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

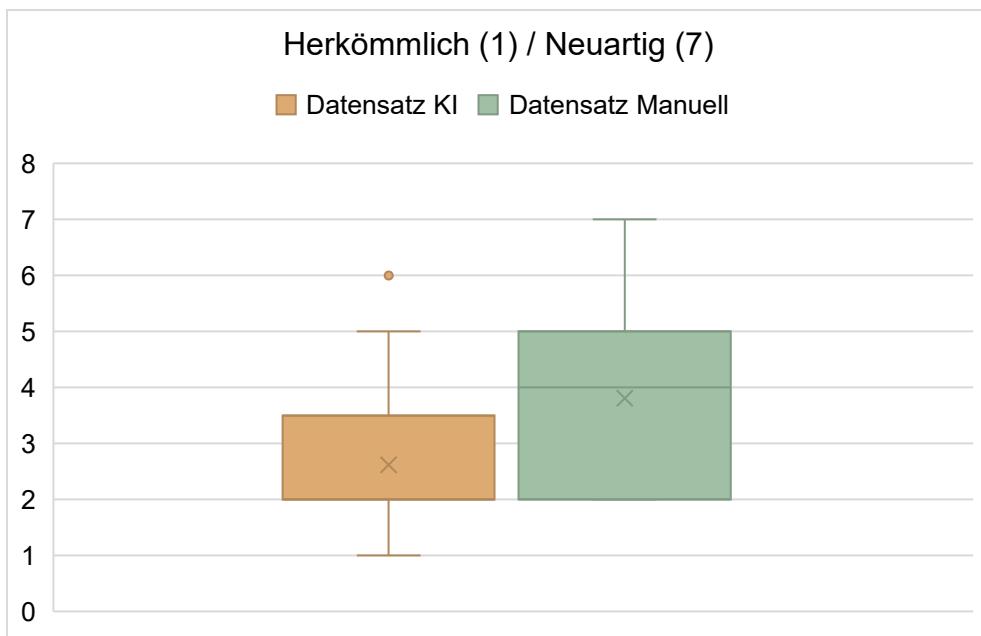


Abbildung 150. Boxplot-Diagramm für die Bewertung der Neuartigkeit

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung der Neuartigkeit zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 2, was einer Bewertung im negativen Bereich entspricht. Der Interquartilsabstand von 1 zeigt eine niedrige Streuung der Bewertungen an, womit die Einschätzung der Neuartigkeit einheitlich ausfiel. Jedoch ist im Diagramm ein Ausreißer bei der Bewertung im oberen positiven Bereich mit dem Wert 6 zu erkennen. Der Mittelwert liegt mit 2,62 ebenfalls im negativen Bereich, weicht jedoch leicht von dem Median ab. Dies weist erneut auf Ausreißer hin.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 4 und einem Interquartilsabstand von 3 darauf hin, dass die Applikation mit einer großen Streuung als neutral bewertet wurde. Hier stimmen die Bewertungen der Nutzer also wenig überein. Auch der Mittelwert von 3,81 befindet sich im mittleren neutralen Bereich und bestätigt den Median.

### Attraktiv (1) / Unattraktiv (7)

Da der Wert  $p = 0,0136$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als attraktiv oder unattraktiv. Der Unterschied ist nicht zufällig entstanden. Der Wert für  $r = 0,38$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

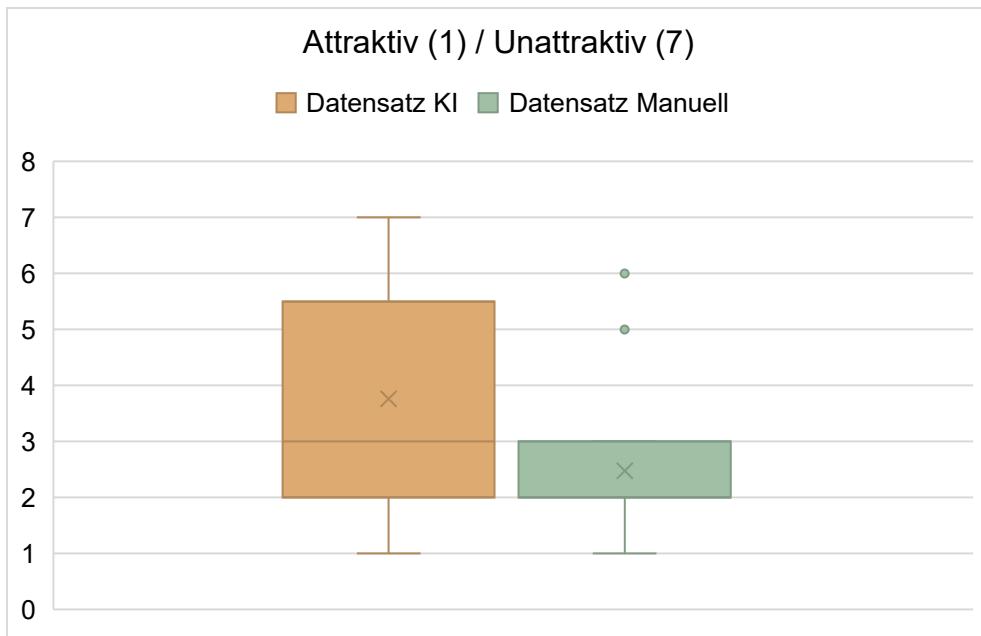


Abbildung 151. Boxplot-Diagramm für die Bewertung der Attraktivität

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung der Attraktivität zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 3, was einer Bewertung im positiven Bereich entspricht. Der Interquartilsabstand von 3 zeigt eine hohe Streuung der Bewertungen an, womit die Einschätzung der Attraktivität nicht einheitlich ausfiel. Der Mittelwert liegt mit 3,76 ebenfalls im positiven Bereich.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 2 und einem Interquartilsabstand von 1 darauf hin, dass die Applikation einheitlich als attraktiv bewertet wurde. Hier stimmen die Bewertungen der Nutzer also überein. Jedoch sind im Diagramm zwei Ausreißer bei der Bewertung im oberen negativen Bereich mit den Werten 5 und 6 zu erkennen. Der Mittelwert von 2,48 befindet sich ebenfalls im unteren positiven Bereich und bestätigt den Median.

## Unpragmatisch (1) / Pragmatisch (7)

Da der Wert  $p = 0,0182$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als nicht pragmatisch oder pragmatisch. Der Unterschied ist nicht zufällig entstanden. Der Wert für  $r = 0,36$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

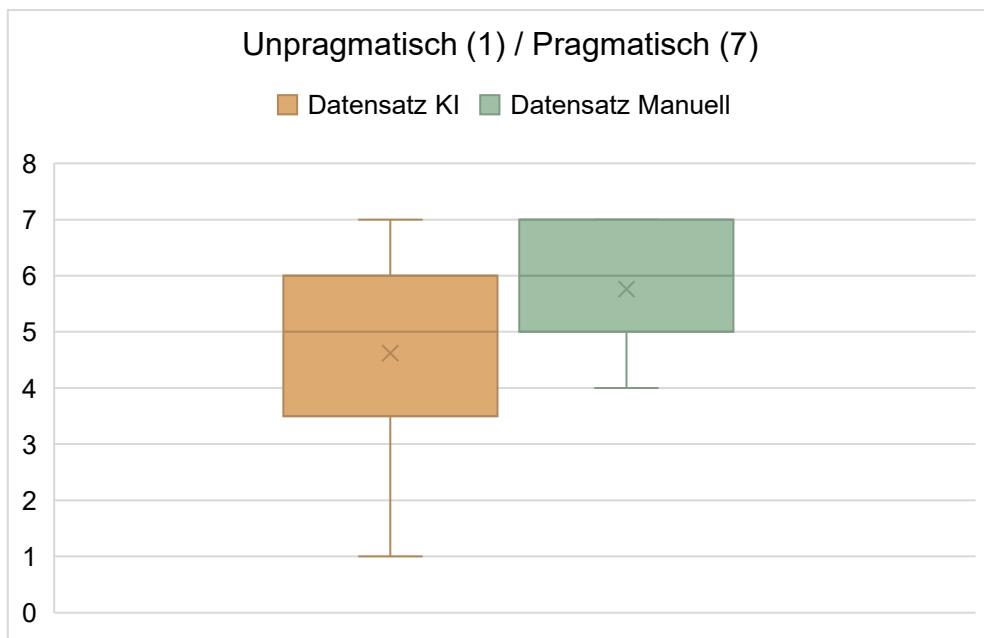


Abbildung 152. Boxplot-Diagramm für die Bewertung der Pragmatik

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung der Pragmatik zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 5, was einer Bewertung im positiven Bereich entspricht. Der Interquartilsabstand von 2 zeigt eine leichte Streuung der Bewertungen an, womit die Einschätzung der Pragmatik nicht ganz einheitlich ausfiel. Der Mittelwert liegt mit 4,62 im positiven Bereich und bestätigt den Median.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 6 und einem Interquartilsabstand von 2 darauf hin, dass die Applikation mit einer leichten Streuung als pragmatisch bewertet wurde. Hier stimmen die Bewertungen der Nutzer also nicht ganz überein. Auch der Mittelwert von 5,76 befindet sich im positiven Bereich und bestätigt den Median.

### Sympathisch (1) / Unsympathisch (7)

Da der Wert  $p = 0,0198$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als sympathisch oder unsympathisch. Der Unterschied ist nicht zufällig entstanden. Der Wert für  $r = 0,36$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

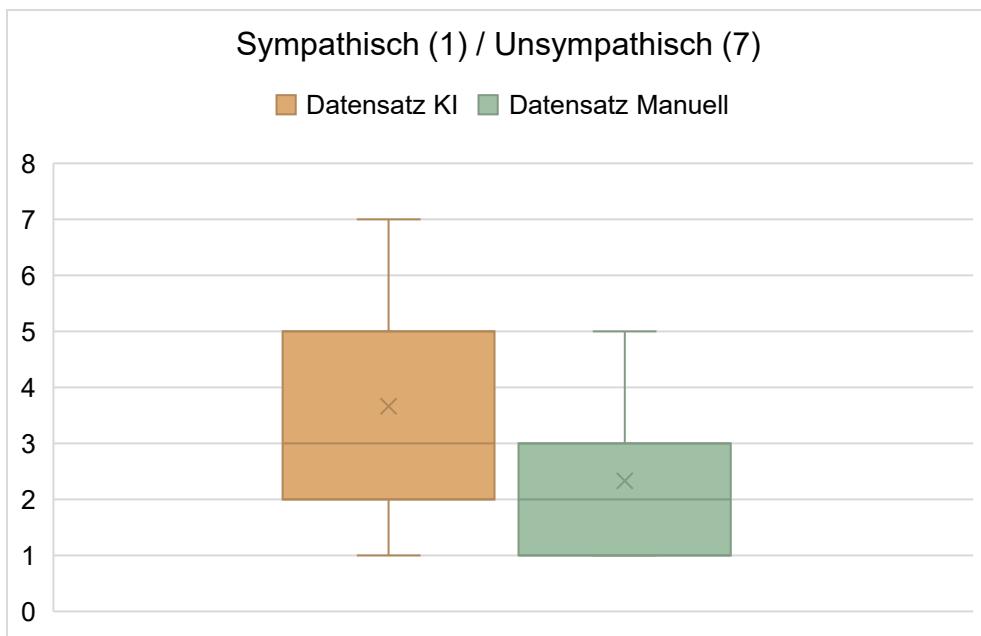


Abbildung 153. Boxplot-Diagramm für die Bewertung der Sympathie

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung der Sympathie zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 3, was einer Bewertung im positiven Bereich entspricht. Der Interquartilsabstand von 3 zeigt eine erhöhte Streuung der Bewertungen an, womit die Einschätzung der Sympathie nicht einheitlich ausfiel. Der Mittelwert liegt mit 3,67 im positiven bis neutralen Bereich und weicht damit leicht vom Median ab. Dies könnte auf Ausreißer in der Bewertung hinweisen.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 2 und einem Interquartilsabstand von 2 darauf hin, dass die Applikation mit einer leichten Streuung als sympathisch bewertet wurde. Hier stimmen die Bewertungen der Nutzer also nur teilweise überein. Auch der Mittelwert von 2,33 befindet sich im positiven Bereich und bestätigt den Median.

### Unberechenbar (1) / Voraussagbar (7)

Da der Wert  $p = 0,0198$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als unberechenbar oder voraussagbar. Der Unterschied ist nicht zufällig entstanden. Der Wert für  $r = 0,36$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

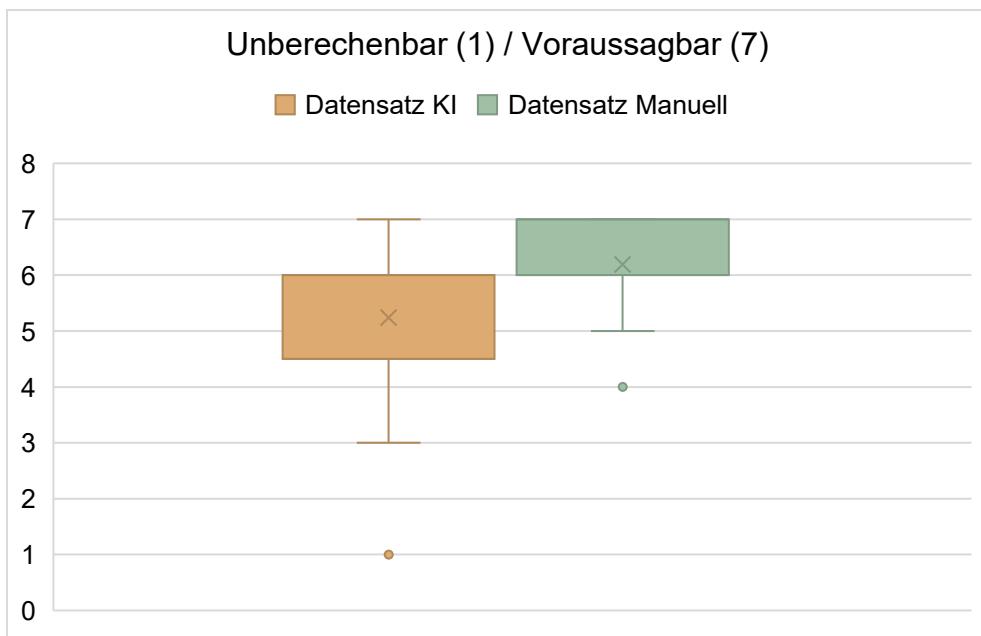


Abbildung 154. Boxplot-Diagramm für die Bewertung der Voraussagbarkeit

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung der Voraussagbarkeit zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 6, was einer Bewertung im positiven Bereich entspricht. Der Interquartilsabstand von 1 zeigt eine niedrige Streuung der Bewertungen an, womit die Einschätzung der Voraussagbarkeit einheitlich ausfiel. Der Mittelwert liegt mit 5,24 ebenfalls im positiven Bereich, weicht jedoch leicht vom Median ab und könnte somit auf leichte Ausreißer in der Bewertung hindeuten.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 6 und einem Interquartilsabstand von 1 darauf hin, dass die Applikation einheitlich als voraussagbar bewertet wurde. Hier stimmen die Bewertungen der Nutzer also überein. Auch der Mittelwert von 6,19 befindet sich im positiven Bereich und bestätigt den Median.

Trotz des durch den Mann-Whitney-U-Test ermittelten statistisch signifikanten Unterschieds in der Einfachheit zwischen der KI-Applikation und der manuellen Applikation für  $p = 0,0198$  und  $r = 0,36$  und dem erkennbaren Unterschied bei den Mittelwerten, ist beim Median kein Unterschied zu erkennen. Dennoch besteht ein Unterschied in der gesamten Verteilung der Werte, wie zum Beispiel in den Extremen oder durch Ausreißer. Ausreißer sind auch in dem Diagramm zu erkennen. Bei der KI-generierten Applikation befinden sich diese im unteren negativen Bereich mit dem Wert 1. Auch die manuell gestaltete Applikation hat einen Ausreißer im mittleren neutralen Bereich mit dem Wert 4.

## Aufgeräumt (1) / Überladen (7)

Da der Wert  $p = 0,0232$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als aufgeräumt oder überladen. Der Unterschied ist nicht zufällig entstanden. Der Wert für  $r = 0,35$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

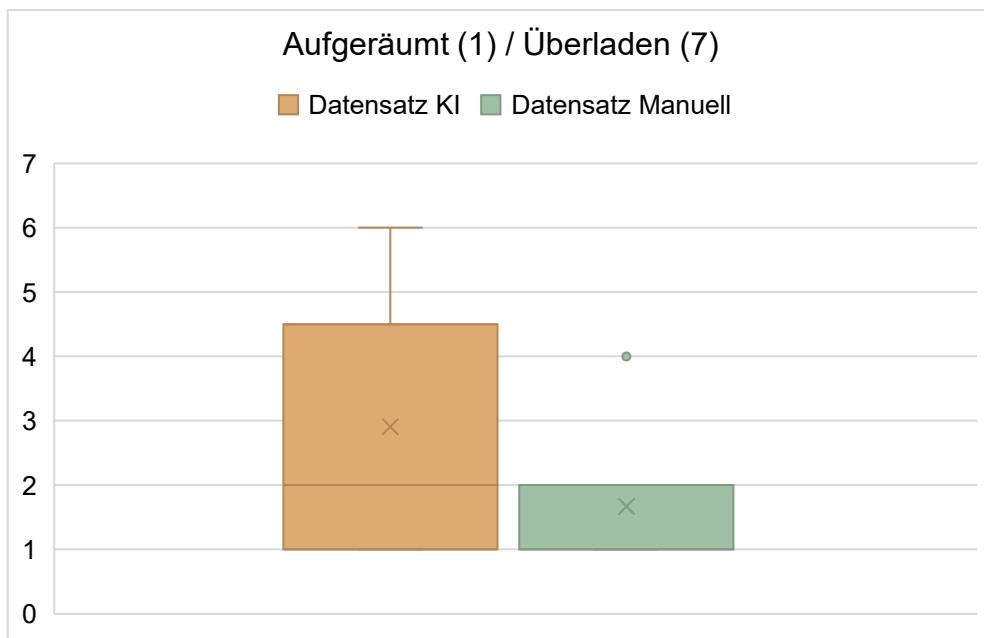


Abbildung 155. Boxplot-Diagramm für die Bewertung der Aufgeräumtheit

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung der Aufgeräumtheit zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 2, was einer Bewertung im positiven Bereich entspricht. Der Interquartilsabstand von 3 zeigt eine erhöhte Streuung der Bewertungen an, womit die Einschätzung der Aufgeräumtheit nicht einheitlich ausfiel. Der Mittelwert liegt mit 2,9 ebenfalls im positiven Bereich, weicht jedoch vom Median ab und könnte auf Ausreißer hinweisen.

Die manuelle Applikation weist mit einem Median von 2 und einem Interquartilsabstand von 1 darauf hin, dass die Applikation als aufgeräumt bewertet wurde. Hier stimmen die Bewertungen der Nutzer also überein. Auch der Mittelwert von 1,67 befindet sich im positiven Bereich und bestätigt den Median.

Trotz des durch den Mann-Whitney-U-Test ermittelten statistisch signifikanten Unterschieds in der Aufgeräumtheit zwischen der KI-Applikation und der manuellen Applikation für  $p = 0,0232$  und  $r = 0,35$  und dem erkennbaren Unterschied bei den Mittelwerten, ist beim Median kein Unterschied zu erkennen. Dennoch besteht ein Unterschied in der gesamten Verteilung der Werte, wie zum Beispiel in den Extremen oder durch Ausreißer. Ausreißer sind für die manuelle Applikation auch in dem Diagramm zu erkennen. Der Ausreißer liegt dabei im mittleren neutralen Bereich mit dem Wert 4.

## Originell (1) / Konventionell (7)

Da der Wert  $p = 0,03$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als originell oder konventionell. Der Unterschied ist dadurch nicht zufällig entstanden. Der Wert für  $r = 0,34$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

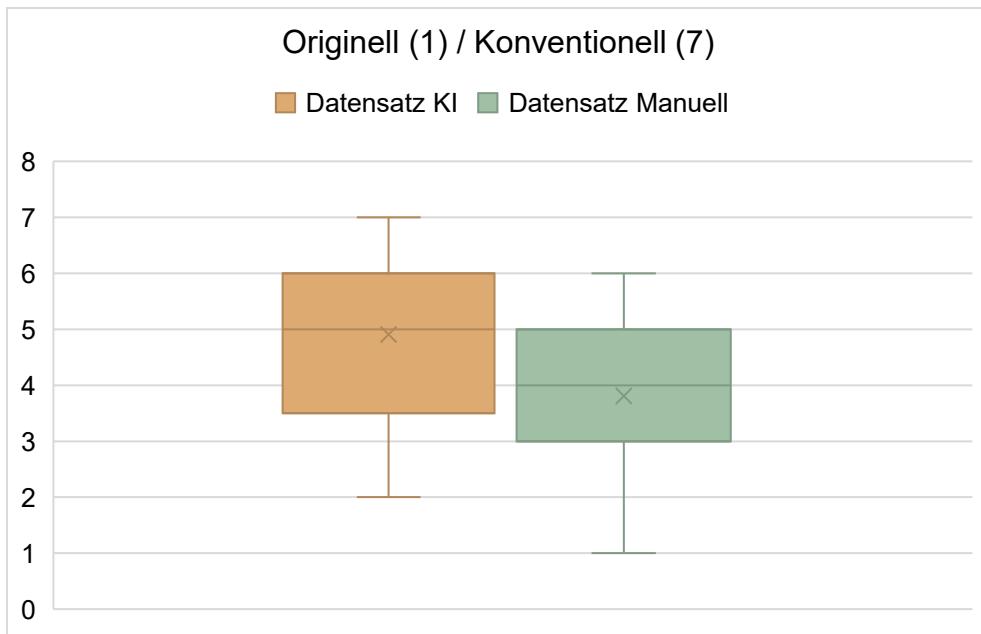


Abbildung 156. Boxplot-Diagramm für die Bewertung der Originalität

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung der Originalität zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 5, was einer Bewertung im negativen Bereich entspricht. Der Interquartilsabstand von 2 zeigt eine leichte Streuung der Bewertungen an, womit die Einschätzung der Originalität nicht ganz einheitlich ausfiel. Der Mittelwert liegt mit 4,9 ebenfalls im negativen Bereich und bestätigt den Median.

Im Vergleich dazu weist die manuelle Applikation mit einem Median von 4 und einem Interquartilsabstand von 2 darauf hin, dass die Applikation mit einer leichten Streuung als neutral bewertet wurde. Hier stimmen die Bewertungen der Nutzer also nicht ganz überein. Auch der Mittelwert von 3,81 befindet sich im neutralen Bereich und bestätigt den Median.

### **Design: Sehr gut (1) / Gar nicht (7)**

Da der Wert  $p = 0,032$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung des Gefallens des Designs als sehr gut oder gar nicht. Der Unterschied ist dadurch nicht zufällig entstanden. Der Wert für  $r = 0,33$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

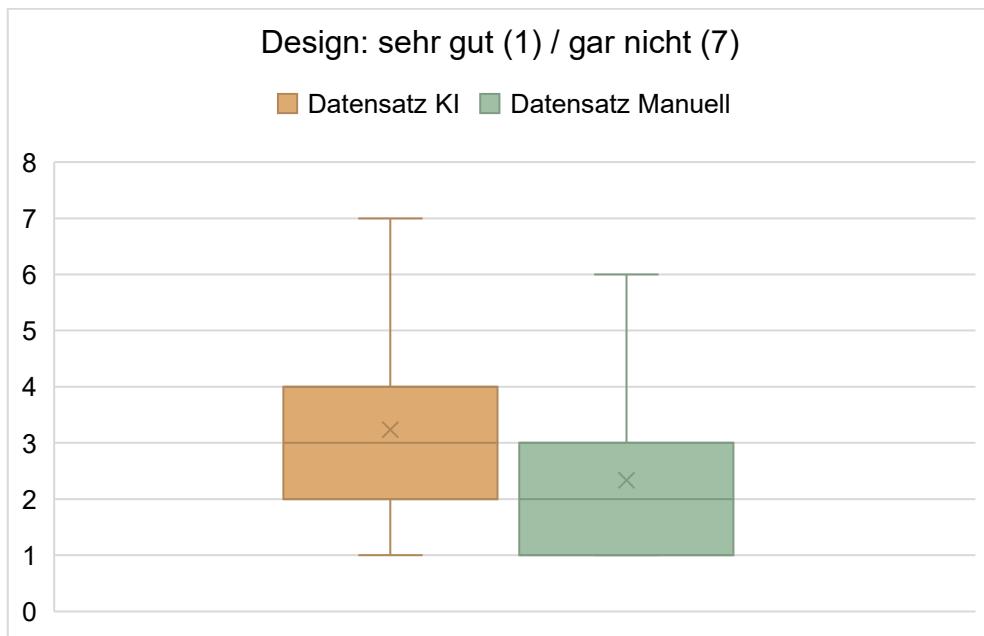


Abbildung 157. Boxplot-Diagramm für die Bewertung des Designs

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung des Designs zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 3, was einer Bewertung im positiven Bereich entspricht. Der Interquartilsabstand von 2 zeigt eine leichte Streuung der Bewertungen an, womit die Einschätzung des Designs nicht ganz einheitlich ausfiel. Der Mittelwert liegt mit 3,24 ebenfalls im positiven Bereich und bestätigt damit den Median.

Die manuelle Applikation weist mit einem Median von 2 und einem Interquartilsabstand von 2 darauf hin, dass das Design der Applikation mit einer leichten Streuung als positiv bewertet wurde. Hier stimmen die Bewertungen der Nutzer also nicht ganz überein. Auch der Mittelwert von 2,33 befindet sich im positiven Bereich und bestätigt den Median.

## **Wertvoll (1) / Minderwertig (7)**

Da der Wert  $p = 0,0394$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen den Bewertungen der beiden Applikationen hinsichtlich ihrer Wahrnehmung als wertvoll oder minderwertig. Der Unterschied ist dadurch nicht zufällig entstanden. Der Wert für  $r = 0,32$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

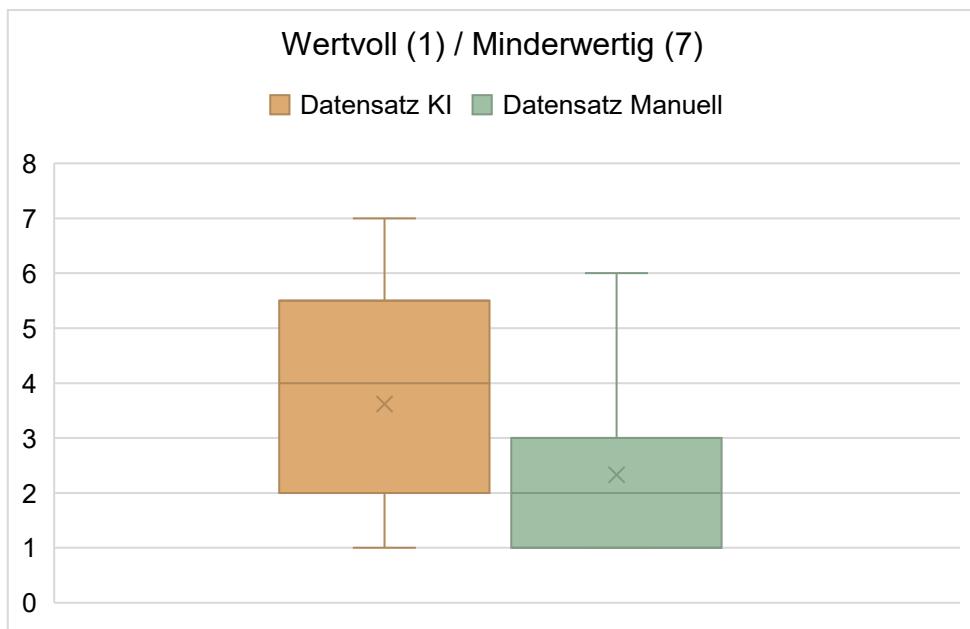


Abbildung 158. Boxplot-Diagramm für die Bewertung des Werts

Am Boxplot-Diagramm zeigt sich ein signifikanter Unterschied bei der Bewertung des Werts zwischen der KI-generierten Applikation und der manuell erstellten Applikation.

Die KI-Applikation hat einen Median von 4, was einer Bewertung im neutralen Bereich entspricht. Der Interquartilsabstand von 3 zeigt eine erhöhte Streuung der Bewertungen an, womit die Einschätzung des Wertes nicht einheitlich ausfiel. Der Mittelwert liegt mit 3,62 ebenfalls annähernd im neutralen Bereich und bestätigt damit den Median.

Die manuelle Applikation weist mit einem Median von 2 und einem Interquartilsabstand von 2 darauf hin, dass die Applikation mit einer leichten Streuung als wertvoll bewertet wurde. Hier stimmen die Bewertungen der Nutzer also nicht ganz überein. Auch der Mittelwert von 2,33 befindet sich im positiven Bereich und bestätigt den Median.

## Aufgabenabschlusszeit

Da der Wert  $p = 0,048$  kleiner als das Signifikanzniveau von 0,05 ist, ist der Unterschied statistisch signifikant. Es besteht also ein Unterschied zwischen der Aufgabenabschlusszeit der beiden Applikationen. Der Unterschied ist demnach nicht zufällig entstanden. Es sollte jedoch beachtet werden, dass der Unterschied nur geringfügig signifikant ist und empfindlich gegenüber kleinen Veränderungen ist. Der Wert für  $r = 0,3$  kann nach Cohen als mittlerer Effekt interpretiert werden und ist damit nicht nur statistisch erkennbar, sondern auch für den Nutzer während der Verwendung wahrnehmbar.

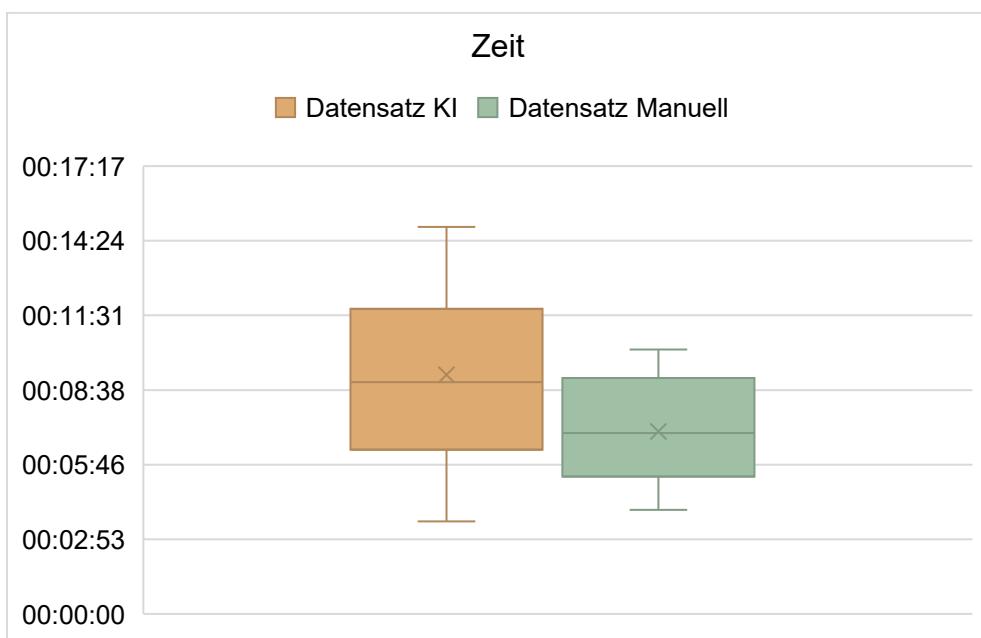


Abbildung 159. Boxplot-Diagramm für die Bewertung der Aufgabenabschlusszeit

Am Boxplot-Diagramm zeigt sich ein leicht signifikanter Unterschied bei den Aufgabenabschlusszeiten der KI-generierten Applikation und der manuell erstellten Applikation.

Mithilfe des Mittelwertes ist zu erkennen, dass die Nutzer im Durchschnitt mehr Zeit für den Abschluss der Aufgaben bei der KI-Applikation benötigten als bei der manuellen Applikation. Der Unterschied beträgt 2 Minuten und 15 Sekunden.

Der Mittelwert der KI-Applikation liegt mit einem Wert von 09:14 leicht über dem Median mit einem Wert von 08:57. Dies könnte daraufhin deuten, dass es Ausreißer in den Zeiten geben könnte, die den Mittelwert erhöht haben.

Auch der Mittelwert der manuellen Applikation liegt mit einem Wert von 07:17 leicht über dem Median mit einem von Wert 06:59. Dies könnte ebenfalls auf Ausreißer in den Zeiten hinweisen.

Die IQR-Werte deuten darauf hin, dass bei der KI-Applikation die Bearbeitungszeiten mit einem Wert von 04:45 etwas stärker gestreut sind als die Zeiten der manuellen Applikation mit einem Wert von 04:10.

## **6.Diskussion der Ergebnisse**

Im Folgenden werden anhand der Ergebnisse der Analysen und der Auswertung der Mitschriften die Forschungsfragen beantwortet.

### **Inwiefern unterscheidet sich die Verwendung einer KI-generierten Applikation zu der Verwendung einer manuell gestalteten Applikation in Bezug auf die Benutzererfahrung?**

Bei den Aspekten aktivierend, kreativ und sicher gab es keinen statistisch signifikanten Unterschied, aber in den Bereichen Innovation, Originalität, Design, Attraktivität, Sympathie und der allgemeinen Bewertung schnitt die KI-Anwendung insgesamt schlechter ab. Die Nutzer empfanden die manuelle Applikation zudem als neuartiger, wertvoller, erfreulicher, anziehender, interessanter, angenehmer und spannender. Die manuell erstellte Applikation wurde damit in fast allen Bereichen der Benutzererfahrung besser beurteilt als die KI-generierte Applikation.

#### **1. Inwiefern unterscheidet sich das Verhalten der Nutzer bei der Verwendung einer KI-generierten Applikation im Vergleich zu der Verwendung einer manuell gestalteten Applikation?**

Bei der Verwendung der KI-Applikation zeigten die Nutzer deutlich häufiger Unsicherheiten, Frustrationen und generierte Reaktionen. Bei der manuell gestalteten Applikation verhielten sich beispielsweise nur vier Personen unsicher wohingegen es bei der KI-generierten Applikation zwanzig Personen waren. Auch weitere Verhaltensweisen wie Zögern und Verzweiflung traten bei der KI-Applikation öfter auf. Im Gegensatz dazu verhielten sich alle einundzwanzig Testpersonen insgesamt sicher bei der Verwendung der manuell erstellten Applikation. Negatives Verhalten wie Zögern, Unsicherheiten und Verwirrungen kamen seltener und nur bei vereinzelten Aufgaben vor. Auffällig ist auch, dass bei der KI-generierten Applikation mehr Nutzer angaben, erst nach einem Lerneffekt gut mit der Website interagieren zu können.

Die manuell erstellte Applikation wirkte also im Vergleich intuitiver und schneller verständlich. Insgesamt zeigten die Nutzer während der Tests bei der manuell gestalteten Applikation mehr positives Verhalten als bei der KI-generierten Applikation.

#### **2. Inwiefern unterscheiden sich die Emotionen der Nutzer bei der Verwendung einer KI-generierten Applikation im Vergleich zu der Verwendung einer manuell gestalteten Applikation?**

Bei der Verwendung der manuell gestalteten Applikation zeigten die Nutzer deutlich häufiger positive Emotionen, wie Freude, Ruhe oder Interesse als bei der KI-generierten Applikation. So empfanden bei der manuell erstellten Applikation einundzwanzig Personen Freude sowie Ruhe wohingegen bei der KI-Applikation nur fünfzehn Testpersonen Freude und niemand Ruhe verspürte. Interesse zeigten bei der manuellen Applikation sieben und bei der KI-Applikation nur ein Nutzer. Negative Emotionen wie Verwirrung, Entsetzen, Wut und Enttäuschung traten jedoch bei der KI-generierten Applikation deutlich häufiger und stärker auf als bei der manuellen Applikation. Verwirrung empfanden beispielsweise achtzehn Personen bei der KI-Applikation und nur sechs bei der manuellen Applikation. Starke Emotionen wie Entsetzen empfanden sechzehn Nutzer und vier fühlten Wut bei der Verwendung der KI-generierten Website. Bei der manuell gestalteten Website fühlten keine Nutzer Entsetzen oder Wut.

Die manuell erstellte Applikation schneidet in Bezug auf die Emotionen deutlich besser ab, da sie häufiger positive Emotionen hervorruft, während negative Emotionen seltener und weniger intensiv auftraten. Im Gegensatz dazu löst die KI-generierte Applikation deutlich häufiger und stärkere negative Emotionen aus und weniger positive Emotionen.

### **Inwiefern unterscheidet sich die Verwendung einer KI-generierten Applikation zu der Verwendung einer manuell gestalteten Applikation in Bezug auf die Benutzerfreundlichkeit?**

Die manuell gestaltete Applikation hat sowohl in der Mitschrift als auch in der Umfrage mehr positives als negatives Feedback erhalten. Das positive Feedback umfasst beispielweise eine gut strukturierte Navigation, eine intuitive und hilfreiche Filterfunktion, einen übersichtlicheren Warenkorb, eine gute Interaktion mit den Elementen, eine gute Auffindbarkeit der Produkte und ein positives, wenn auch etwas konservatives Design. Es gab auch vereinzelt kleinere Kritikpunkte von wenigen Nutzern, wie das Finden des Matcha-Sets, die Beendigung des Bestellprozesses oder die Designentscheidungen.

Die KI-generierte Applikation erhielt jedoch sowohl in der Mitschrift als auch in der Umfrage mehr negatives als positives Feedback. Das negative Feedback umfasste beispielsweise die eingeschränkten Filtermöglichkeiten, viele Designprobleme, wie unpassende Farben und eine langweilige und lustlose Gestaltung, eine oft unübersichtliche Navigation, eine fehlende Suchfunktion, fehlerhafte Buttons auf der Startseite sowie das erschwerte Auffinden der Produkte. Das Design und die Navigation wurde zwar von ein paar Nutzern als positiv bewertet, jedoch überwiegte die Anzahl der Personen, die dies negativ bewerteten. Weiteres positives Feedback waren die Möglichkeit einer Lernkurve, ein gewohnter Warenkorb, das Auffinden der vorhandenen Filter und die Interaktion mit den Elementen.

Insgesamt erhielt die manuell gestaltete Applikation ein positiveres Feedback als die KI-generierte Applikation.

#### **1. Inwiefern unterscheidet sich die Effizienz bei der Verwendung einer KI-generierten Applikation im Vergleich zu der Verwendung einer manuell gestalteten Applikation?**

Bezüglich der Effizienz ergab sich kein statistisch signifikanter Unterschied in den Bereichen Schnelligkeit und Erlernbarkeit. Allerdings benötigten die Nutzer bei der KI-Applikation mehr Zeit für den Abschluss der Aufgaben, im Durchschnitt 9:14 Minuten gegenüber einem Durchschnitt von 6:59 Minuten bei der manuellen Applikation. Die manuell erstellte Applikation wurde zudem als einfacher, effizienter, besser navigierbar und besser sortiert wahrgenommen.

#### **2. Inwiefern unterscheidet sich die Effektivität bei der Verwendung einer KI-generierten Applikation im Vergleich zu der Verwendung einer manuell gestalteten Applikation?**

Auch in der Effektivität ergaben sich statistisch signifikante Unterschiede. Die manuelle Applikation wurde als übersichtlicher, verständlicher, effizienter und voraussagbarer beurteilt. Im Vergleich hatte die KI-Applikation hier ebenfalls schlechtere Werte. Bezüglich der Erlernbarkeit ergab sich jedoch kein statistisch signifikanter Unterschied.

## Abschlussquote

Bei beiden Applikationen konnten zwei der einundzwanzig Nutzer nicht alle Aufgaben abschließen. Jedoch wurden bei der KI-generierten Applikation insgesamt drei Aufgaben nicht abgeschlossen wohingegen es bei der manuell gestalteten Applikation nur zwei waren. Die Gründe für das nicht Abschließen der Aufgaben bei der KI-Applikation waren das Übersehen der Banner, schwer auffindbare Tees ohne Filter und das falsche Interpretieren des Bio-Filters. Bei der manuell gestalteten Applikation bestanden die Fehler aus der Unklarheit über den beendeten Abschluss des Bestellprozesses und der Auswahl eines falschen Schwarzees. Ein weiterer Unterschied besteht auch in der Art der Fehler. Bei der KI-generierten Applikation entstanden die Fehler hauptsächlich aufgrund von Interface-Schwächen, da zum einen die Banner übersehen wurden und damit wichtige Informationen, die zum Auffinden der Filter benötigt werden, nicht sichtbar genug waren. Zum anderen war die Filterlogik nicht deutlich genug, wodurch die Nutzer nicht verstanden, welche Tees der Filter zeigte. Die Probleme bestanden also eher in der Navigation, der Übersichtlichkeit und der Verständlichkeit der Anwendung. Die Fehler bei der manuell gestalteten Applikation entstanden eher durch individuelle Missverständnisse, da die Testpersonen nicht realisierten, dass der Bestellprozess abgeschlossen war, obwohl er es war, und ein falscher Tee ausgewählt wurde wegen fehlender weiterer Filteranwendung, aber dennoch generell korrekter Bedienung. Die Probleme bestanden also eher in kleinen Verständnisfragen und nicht in Interface- oder Navigations-Schwächen der Anwendung.

Obwohl bei beiden Applikationen gleich viele Aufgaben nicht erfolgreich abgeschlossen wurden, schneidet die manuell gestaltete Applikation etwas besser ab. Die Fehler bei der manuellen Applikation waren weniger schwerwiegend und entstanden eher auf Grund von Verständnisproblemen einzelner Nutzer. Die KI-Applikation zeigt hingegen grundlegende Usability-Schwächen wie eine schlechte Sichtbarkeit von Informationen und eine unklare Filterlogik, die zu funktionalen Problemen führen und die Bedienung erschweren.

### **3. Inwiefern unterscheidet sich die Zufriedenheit der Nutzer bei der Verwendung einer KI-generierten Applikation im Vergleich zu der Verwendung einer manuell gestalteten Applikation?**

Auch in der Nutzerzufriedenheit konnten statistisch signifikante Unterschiede festgestellt werden. Die manuell gestaltete App wurde als unterstützender, erwartungskonformer, pragmatischer und aufgeräumter bewertet. Die KI-Applikation schneidet im Vergleich in allen Kategorien schlechter ab.

### **4. Inwiefern unterscheidet sich das Gefallen der visuellen Erscheinung bei der Verwendung einer KI-generierten Applikation im Vergleich zu der Verwendung einer manuell gestalteten Applikation?**

Die KI erstellte Applikation wird bezüglich des Designs von den Nutzern mit einem statistisch signifikanten Unterschied schlechter bewertet als die manuell erstellte Applikation.

Das Design der manuell gestalteten Applikation wurde von den Nutzern hauptsächlich positiv bewertet. Die schöne sowie angenehme Farbwahl, die aufgeräumte und übersichtliche Gestaltung sowie die moderne, jedoch eher konservative Designsprache wurden als unterstützend für ein schnelles Auffinden von Informationen und Produkten angesehen. Jedoch gab es auch negatives

Feedback. Zum einen, dass das Design überarbeitet werden sollte, insbesondere hinsichtlich einer lebendigeren Farbgestaltung für mehr „Pepp“ und Tiefe. Zum anderen wurde angemerkt, dass die Hintergrundfarbe in eine auffälligere Farbe geändert werden sollte. Insgesamt übertraf jedoch die Anzahl des positiven Feedbacks das negative.

Das Design der KI-generierten Applikation wurde als ansprechend, modern, generisch und zugleich schlicht mit schönen Farben bewertet. Es ergab sich jedoch auch negatives Feedback. Insbesondere die Verwendung der blauen Schrift auf gelbem Hintergrund, das Layout der Banner, das Logo sowie der Footer und dessen Inhalte wurden kritisiert. Weiterhin wurde der Warenkorb-Button von mehreren Nutzern als zu klein empfunden und die Website allgemein hatte zu viele visuelle Effekte. Dadurch übertraf das negative Feedback das positive.

Das Feedback der Nutzer bestätigt die Auswertung der Analyse der Umfrage, da auch hier deutlich wird, dass das Design der manuell gestalteten Applikation als besser empfunden wird als das der KI-generierten Applikation.

## 7. Fazit

### 7.1 Zusammenfassung

Diese Arbeit untersuchte, welche Unterschiede es in der Verwendung einer KI-generierten Applikation zu der Verwendung einer manuell gestalteten Applikation bezüglich der Benutzererfahrung und der Benutzerfreundlichkeit gibt. Ziel dieser Arbeit war es herauszufinden, ob KI-generierte Websites tatsächlich eine überzeugende Alternative für Endbenutzer darstellen.

Es ergaben sich dabei deutliche Unterschiede in der Benutzererfahrung sowie in dem Verhalten und den Emotionen. Dabei ergab sich meist ein positiveres Feedback für die manuell gestaltete Applikation. Lediglich in den Bereichen aktivierend, kreativ oder sicher konnten keine Unterschiede zwischen den beiden Applikationen erkannt werden.

Auch in der Benutzerfreundlichkeit ergaben sich Unterschiede sowohl in den Umfragen als auch im Feedback während der Tests. Dabei erhielt die manuelle Applikation mehr positive Bewertungen als die KI-generierte Applikation. Ähnlich verhielt es sich auch bei der Bewertung der visuellen Erscheinung. Dabei wurde das Design der manuellen Applikation als angenehm, übersichtlich und funktional empfunden. Kritisiert wurden nur kleinere Designentscheidungen. Die Bewertung der KI-Applikation fiel in diesen Bereichen ebenfalls schlechter aus und es wurde mehr kritisches Feedback bezüglich der Designentscheidungen genannt. Zudem ergaben sich auch Unterschiede in dem Bereich Effizienz und Effektivität, wobei die manuelle Applikation, bis auf den Bereichen Schnelligkeit und Erlernbarkeit, positiver bewertet wurde.

Insgesamt lässt sich also sagen, dass die manuell gestaltete Applikation meistens besser bewertet wird als die KI-generierte Applikation. Jedoch bedeutet das nicht zwingend, dass die KI-Anwendung schlecht ist. Werden die Mittelwerte beider Applikationen aus den Umfragen getrennt betrachtet fällt auf, dass die der KI-Applikation meist im neutralen oder positiven Bereich liegen. Demnach ist die Applikation laut

den Nutzern eine durchaus zufriedenstellende Website, die in der Praxis funktionieren kann, jedoch aktuell noch nicht mit einer vom Webdesigner entwickelten Website mithalten kann.

## 7.2 Zukunftsausblick

Abschließend ist festzuhalten, dass sich die künstliche Intelligenz ständig weiterentwickelt und sicherlich in den nächsten Jahren noch große Fortschritte machen wird, auch im Bereich Webdesign. Vereinzelte Anpassungen werden auch notwendig sein, da die Erstellung der Website aktuell für einen Laien noch herausfordernd und frustrierend sein kann, zum Beispiel beim Erstellen von Produktseiten, beim Festlegen von Versandoptionen oder beim Schreiben von klaren Anweisungen an die KI über den Chat. Insgesamt sollte der Erstellungsprozess für einen Laien noch weiter vereinfacht und verbessert werden. An dieser Stelle sollte zusätzlich angemerkt werden, dass die Erstellung der KI-generierten Website aktuell noch lange dauert und der Ersteller bereit sein muss diese Zeit aufzuwenden. Es bleibt abzuwarten, ob größere Firmen bereit sind, Mitarbeiter dafür über einen längeren Zeitraum intern zu beschäftigen oder ob sie lieber weiterhin auf Webdesigner und Webagenturen setzen, die auch eine langfristige Betreuung anbieten. Für kleinere Startups jedoch, die schnell eine einfache und kostengünstige Alternative benötigen, könnten KI-generierte Applikationen auch jetzt schon eine gute Alternative darstellen. Immerhin war das Endergebnis der KI-Applikation durchaus zufriedenstellend, da die Bewertungen sich meist im neutralen und positiven Bereich befanden, jedoch gibt es noch einen eindeutigen Verbesserungsbedarf bis zu einer benutzerfreundlichen Website, die mit den Ergebnissen einer manuell gestalteten Website mithalten kann.

Wenn es der KI jedoch gelingt, die noch bestehenden Schwächen zu beheben, könnte sie in Zukunft langfristig eine gute Alternative für die Nutzer darstellen.

## 7.3 Bewertung der Arbeit

Die Zielsetzung der Arbeit konnte im Wesentlichen erfüllt werden. Die Forschungsfrage wurde durch eine Kombination aus qualitativer Analyse und empirischen Daten beantwortet, auch wenn die Tiefe der Auswertung durch eine geringe Umfrageanzahl eingeschränkt war. Die Anzahl der Testteilnehmer reichte zwar grundsätzlich aus, um eine valide Datenanalyse durchzuführen, jedoch hätte eine größere Anzahl als einundzwanzig Teilnehmer pro Applikation zu weiteren tiefergehenden interessanten Ergebnissen geführt. Die Usability-Tests stellten dadurch die größte Herausforderung dar, da es schwierig war, ausreichend Teilnehmer in einem ursprünglich geplanten Zeitraum von einem Monat zu erreichen. Nachdem der Testzeitraum verlängert wurde, konnten jedoch auch durch Unterstützung von Prof. Dr.-Ing. Sebastian Stadler ausreichend Teilnehmer gefunden werden.

Weiterhin war der Umfang der Arbeit mit der zeitaufwendigen Erstellung der manuellen Applikation und der großen Anzahl an auszuwertenden Daten des User Experience Questionnaire sehr groß. Letzteres ermöglichte zwar eine ausführliche Bewertung der Applikationen, erzeugte aber auch viele zeitaufwendige Rechenarbeiten, deren Ergebnisse anschließend interpretiert werden mussten. Demnach ergab sich eine umfassende Bachelorarbeit.

Zuletzt gilt es zu erwähnen, dass es bei dem Datensatz der Schnelligkeit für die manuelle Applikation einen negativen Ausreißer mit dem Wert 7 gab, der den statistisch signifikanten Unterschied negativ

beeinflusst haben könnte, da der Unterschied nur knapp statistisch nicht signifikant und sensibel gegenüber Veränderungen ist. Dieser Verdacht bestätigt sich, da die gleiche Person die Applikation stets sehr positiv bewertete. Diese Bewertung könnte versehentlich entstanden sein, da die Begriffe in zufälliger Reihenfolge angeordnet wurden, sodass bei der Hälfte der Fragen der positive Begriff zuerst steht und bei der anderen Hälfte der negative Begriff. Viele Testteilnehmer kritisierten dies nach dem Test auch und gaben an sie hätten oft fast eine falsche Bewertung abgegeben. In zukünftigen Arbeiten mit den User Experience Questionnaire sollte darüber nachgedacht werden, ob auf eine abwechselnde Reihenfolge verzichtet werden kann.

## 8. Verwendete Tools

Für die Durchführung und Entwicklung des Projektes waren verschiedene Tools notwendig.

- Notion als Tool für das Projektmanagement
- Miro zur Darstellung und Bearbeitung der Methoden aus dem human-centered Design Prozess
- Figma zur Entwicklung des Low-Fidelity- und der High-Fidelity-Prototypen
- KI-Website Builder von WIX<sup>16</sup> zur Erstellung der KI-Applikation
- Statistik Rechner von DATAtab<sup>17</sup> zur Überprüfung der berechneten Mann-Whitney-U-Testergebnisse
- Excel zur Darstellung und Analyse der Umfragewerte sowie der Erstellung der Diagramme
- Pixabay<sup>18</sup> für die Bilder manuell gestalteten und die KI-generierte Applikation
- Google Maps<sup>19</sup> für das Bild des Standortes für die manuelle Applikation

---

<sup>16</sup> [https://de.wix.com/ai-website-builder?utm\\_source=google&utm\\_medium=cpc&utm\\_campaign=22085671427^173305799419^search%20-%20wix&experiment\\_id=wix%20ki^e^728036052423^&gad\\_source=1&gclid=Cj0KCQjwoNzABhDbARI-sALFY8VNvW1xbTcrU0b5k0g32kMqco5e18LTWYijj4GFldvP5l2Gdp1lhJYaAh5oEALw\\_wcB](https://de.wix.com/ai-website-builder?utm_source=google&utm_medium=cpc&utm_campaign=22085671427^173305799419^search%20-%20wix&experiment_id=wix%20ki^e^728036052423^&gad_source=1&gclid=Cj0KCQjwoNzABhDbARI-sALFY8VNvW1xbTcrU0b5k0g32kMqco5e18LTWYijj4GFldvP5l2Gdp1lhJYaAh5oEALw_wcB)

<sup>17</sup> <https://datatab.de/statistik-rechner/hypothesentest/mann-whitney-u-test?example=mann-whitney-u-test>

<sup>18</sup> <https://pixabay.com/>

<sup>19</sup> [https://www.google.com/maps/place/Ob.+Karlstra%C3%9Fe,+91054+Erlangen/@49.5965495,11.0061401,571m/data=!3m2!1e3!4b1!4m6!3m5!1s0x47a1f8dfda447fe7:0xcd177eca70369d21!8m2!3d49.5965461!4d11.008715!16s%2Fg%2F1tz94rk0?entry=ttu&g\\_ep=EgoyMDI1MDQzMCA4wIKXMDSoASAFQAw%3D%3D](https://www.google.com/maps/place/Ob.+Karlstra%C3%9Fe,+91054+Erlangen/@49.5965495,11.0061401,571m/data=!3m2!1e3!4b1!4m6!3m5!1s0x47a1f8dfda447fe7:0xcd177eca70369d21!8m2!3d49.5965461!4d11.008715!16s%2Fg%2F1tz94rk0?entry=ttu&g_ep=EgoyMDI1MDQzMCA4wIKXMDSoASAFQAw%3D%3D)

## 9. Referenzen

Agente (2019). Grafik: Usability Heuristics for User Interface Design. Abgerufen am 03.04.2025, von <https://www.epicpxls.com/items/set-of-free-posters-for-user-interface-design>

Behnke, J., & Behnke, N. (2006). *Grundlagen der statistischen Datenanalyse: Eine Einführung für Politikwissenschaftler* (1. Aufl.). VS Verlag für Sozialwissenschaften.

Beranek, A. (2021). Auch eine Frage der Menschenrechte: Was ist künstliche Intelligenz? *Politikum*, 7(1), 4-11. <https://doi.org/10.46499/1608>

Berg, A. (2021, 21. April). *Digitalisierung der Wirtschaft*. Bitkom. Abgerufen am 31.07.2024, von [https://www.bitkom.org/sites/default/files/2021-04/bitkom-charts-kunstliche-intelligenz-21-04-2021\\_final.pdf](https://www.bitkom.org/sites/default/files/2021-04/bitkom-charts-kunstliche-intelligenz-21-04-2021_final.pdf)

Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (7. Aufl.). Springer Berlin Heidelberg.

Büchel, J., Dr. Demary, V., Dr. Goecke, H., Kohlisch, E., Dr. Koppel, O., Dr. Mertens, A., Dr. Rusche, C., Dr. Scheufen, M., & Wendt, J. (2021, August). *www.bvdw.org KI-Monitor 2021 Status quo der Künstlichen Intelligenz in Deutschland Gutachten*. Institut der deutschen Wirtschaft. Abgerufen am 31.07.2024, von [https://www.iwkoeln.de/fileadmin/user\\_upload/Studien/Gutachten/PDF/2021/KI\\_Monitor\\_Bericht\\_2021.pdf](https://www.iwkoeln.de/fileadmin/user_upload/Studien/Gutachten/PDF/2021/KI_Monitor_Bericht_2021.pdf)

*Bürgerliches Gesetzbuch (BGB) § 106 Beschränkte Geschäftsfähigkeit Minderjähriger*. (n.d.). Gesetze im Internet. Abgerufen am 16.01.2025, von [https://www.gesetze-im-internet.de/bgb/\\_106.html](https://www.gesetze-im-internet.de/bgb/_106.html)

*Bürgerliches Gesetzbuch (BGB) § 107 Einwilligung des gesetzlichen Vertreters*. (n.d.). Gesetze im Internet. Abgerufen am 16.01.2025, von [https://www.gesetze-im-internet.de/bgb/\\_107.html](https://www.gesetze-im-internet.de/bgb/_107.html)

Burghardt, M., Heckner, M., Kattenbeck, M., Schneidermeier, T., & Wolff, C. (2011). Design Thinking = Human-Centered Design? In Eibl, M. *Workshop-Proceedings der Tagung Mensch & Computer 2011: überMEDIEN-ÜBERmorgen* (S. 363-368). Univ.-Verlag.

Burkhardt, N. (2018). *Das große Handbuch Innovation: 555 Methoden und Instrumente für mehr Kreativität und Innovation im Unternehmen* (B. Aerssen & C. Buchholz, Eds.). Vahlen.

Chapter 4. Miller's Law. (n.d.). O'Reilly. Abgerufen am 09.04.2025, von <https://www.oreilly.com/library/view/laws-of-ux/9781492055303/ch04.html>

Chlebek, P. (2011). *Praxis der User Interface-Entwicklung: Informationsstrukturen, Designpatterns, Vorgehensmuster* (1. Aufl.). Vieweg+Teubner Verlag.

Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Lawrence Erlbaum Associates.

*Digitale Technologien - Künstliche Intelligenz.* (n.d.). digitale-technologien.de. Abgerufen am 31.07.2024, von <https://www.digitale-technologien.de/DT/Navigation/DE/Themen/KuenstlicheIntelligenz/KuenstlicheIntelligenz.html>

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, 141(1), 2–18. [doi:10.1037/a0024338](https://doi.org/10.1037/a0024338). Abgerufen am 25.04.2025, von [https://www.researchgate.net/publication/51554230\\_Effect\\_Size\\_Estimates\\_Current\\_Use\\_Calculations\\_and\\_Interpretation](https://www.researchgate.net/publication/51554230_Effect_Size_Estimates_Current_Use_Calculations_and_Interpretation)

Geis, T., & Tesch, G. (2019). *Basiswissen Usability und User Experience* (1. Aufl.) dpunkt.verlag.

*Geschäfts- und Deliktsfähigkeit.* (n.d.). Rechtskunde Online. Abgerufen am 16.01.2025, von <https://www.uni-potsdam.de/de/rechtskunde-online/rechtsgebiete/zivilrecht/geschaefts-und-deliktsfaehigkeit>

Hahn, M. (2020). *Webdesign: das Handbuch zur Webgestaltung* (3. Aufl.). Rheinwerk Verlag.

Heimgärtner, R. (2017). *Interkulturelles User Interface Design: Von der Idee zum erfolgreichen Produkt*. Springer Berlin Heidelberg.

Hinderks, A., Schrepp, M., & Thomaschewski, J. (n.d.). *User Experience Questionnaire*. UEQ. Abgerufen am 03.01.2025, von <https://www.ueq-online.org/>

Hoffmann, S. (2023). *Digitales Produktmanagement: Methoden - Instrumente - Praxisbeispiele* (2. Aufl.). Springer Fachmedien Wiesbaden GmbH.

*Homepage-Baukasten » Eigene Website mit IONOS*. (n.d.). IONOS. Abgerufen am 02.10.2024, von <https://www.ionos.de/websites/homepage-baukasten>

Homrighausen, H. (2020). Klett Ich kann ... Mathe - Wahrscheinlichkeitsrechnung 7./8. Klasse: Matematik Schritt für Schritt verstehen. Deutschland: Klett Lerntraining bei PONS.

Jacobsen, J., & Meyer, L. (2022). *Praxisbuch Usability und UX: was alle wissen sollten, die Websites und Apps entwickeln*. Rheinwerk Computing.

Kaplan, K. (2020, 14. Juni). *Approaches to Journey Mapping: 2 Critical Decisions To Make Before You Begin*. Nielsen Norman Group. Abgerufen am 04.04.2025, von <https://www.nngroup.com/articles/journey-mapping-approaches/>

Kauer-Franz, M., & Franz, B. (2022). Usability und User Experience Design: das umfassende Handbuch. Rheinwerk Verlag.

*KI Website erstellen | Kostenlose KI Website in wenigen Minuten*. (n.d.). Wix. Abgerufen am 30.01.2025, von <https://de.wix.com/ai-website-builder>

*Künstliche Intelligenz*. (n.d.). Statista. Abgerufen am 31.07.2024, von <https://de.statista.com/statistik/kategorien/kategorie/15/themen/2604/branche/kuenstliche-intelligenz/#statistic4>

*Künstliche Intelligenz - Perspektive der deutschen Gesellschaft*. (2024, Juli 16). Statista. Abgerufen am 31.07.2024, von <https://de.statista.com/themen/9434/ki-in-der-deutschen-gesellschaft/#topicOverview>

*Künstliche Intelligenz - Perspektive der deutschen Wirtschaft*. (2024, July 9). Statista. Abgerufen am 31.07.2024, von <https://de.statista.com/themen/9400/ki-in-der-deutschen-wirtschaft/#topicOverview>

Lewrick, M. (2020). *Das Design Thinking Toolbook: Die besten Werkzeuge & Methoden* (M. Lewrick, P. Link, & L. Leifer, Eds.). Versus Verlag.

Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 152–158). Association for Computing Machinery. 10.1145/191666.191729

Nielsen, J. (2021, 16. April). *The Relationship Between Artificial Intelligence and User Experience* (Video). Nielsen Norman Group. Abgerufen am 31.07.2024, von <https://www.nngroup.com/videos/relationship-ai-ux/>

Online Statistik Rechner. (n.d.). DATAtab. Abgerufen am 05.04.2025, von <https://datatab.de/statistik-rechner/hypothesentest/mann-whitney-u-test?example=mann-whitney-u-test>

Otte, R. (2021). *Allgemeinbildung Künstliche Intelligenz: Risiko und Chance für dummies* (1. Aufl.). Wiley-VCH.

Paneru, B., Paneru, B., Poudyal, R., & Shah, K. B. (2024, 1. Juni). Exploring the Nexus of User Interface (UI) and User Experience (UX) in the Context of Emerging Trends and Customer Experience, Human Computer Interaction, Applications of Artificial Intelligence. *International Journal of Informatics, Information System and Computer Engineering*, 5(1), 102-113. <https://doi.org/10.34010/injiiscom.v5i1.12488>

Quirmbach, S. M. (2012). *Suchmaschinen: User Experience, Usability und nutzerzentrierte Website-Gestaltung*. Springer Berlin Heidelberg.

Rasch, B., Friese, M., Hofmann, W., & Naumann, E. (2014). *Quantitative Methoden 2: Einführung in die Statistik für Psychologen und Sozialwissenschaftler* (4. Aufl.). Springer Berlin Heidelberg.

Rustambek, M. (2023, 5. Juli). THE ROLE OF ARTIFICIAL INTELLIGENCE IN WEB SITES. *Proceedings of International Conference on Scientific Research in Natural and Social Sciences*, 2(7), 101–107.

Schäfer, T. (2016). *Methodenlehre und Statistik: Einführung in Datenerhebung, deskriptive Statistik und Inferenzstatistik*. Springer Fachmedien Wiesbaden.

Schmidt, F. (2021). Sexismus und Rassismus im Code: Diskriminierung oder Gerechtigkeit durch Algorithmen? *Politikum*, 7(1), 4-11. <https://doi.org/10.46499/1608>

Schrepp, M. (2023, 12. September). *User Experience Questionnaire Handbook - All you need to know to apply the UEQ successfully in your projects* (11) [PDF]. UEQ. Abgerufen am 07.01.2025, von <https://www.ueq-online.org/Material/Handbook.pdf>

Santrel Media. (2024, 26. Oktober). *How To Build A Website in 16 Minutes (Wix AI Builder Tutorial 2024)* [Video]. YouTube. <https://www.youtube.com/watch?v=H3zmx4OljLY>

Sponheim, C., & Brown, M. (2024, 12. April). *Status Update: AI UX-Design Tools Are Not Ready for Primetime*. Nielsen Norman Group. Abgerufen am 31.07.2024, von <https://www.nngroup.com/articles/ai-design-tools-not-ready/>

Thesmann, S. (2009). *Einführung in das Design multimedialer Webanwendungen* (1. Aufl.). Vie weg+Teubner Verlag.

Tokarski, K. O. (2008). *Ethik und Entrepreneurship: Eine theoretische sowie empirische Analyse junger Unternehmen im Rahmen einer Unternehmensethikforschung* (1. Aufl.). Gabler Verlag.

Uebernickel, F., Brenner, W., Pukall, B., Naef, T., & Schindholzer, B. (2015). *Design Thinking: Das Handbuch* (1. Aufl.). Frankfurter Allgemeine Buch.

Uni Hamburg. (n.d.). *Verteilungstabellen*. Abgerufen am 23.04.2025, von <https://www.bwl.uni-hamburg.de/matstat/studium/sose2015/regressionsmodelle/verteilungstabellen.pdf>

*Was ist künstliche Intelligenz und wie wird sie genutzt?* (2023, 20. Juni). Was ist künstliche Intelligenz und wie wird sie genutzt? | Themen | Europäisches Parlament. Abgerufen am 31.07.2024, von <https://www.europarl.europa.eu/topics/de/article/20200827STO85804/was-ist-kunstliche-intelligenz-und-wie-wird-sie-genutzt>

Yablonski, J. (2024). *Laws of UX: 10 praktische Grundprinzipien für intuitives, menschenzentriertes UX-Design* (I. Kommer & C. Kommer, Trans.; 2. Aufl.). Dpunkt.Verlag GmbH.

Zanker, C., Roth, I., & Hoppe, M. (2019, November). *ver.di-Innovationsbarometer 2019 - Künstliche Intelligenz*. input-consulting. Abgerufen am 07.04.2025, von [https://www.input-consulting.de/files/inpcon-DATA/download/2019\\_verdi-Innovationsbarometer2019-KI\\_INPUTConsulting.pdf](https://www.input-consulting.de/files/inpcon-DATA/download/2019_verdi-Innovationsbarometer2019-KI_INPUTConsulting.pdf)

# Anhangsverzeichnis

Anhang 1: Fiktive Auftragsstellung .....	171
Anhang 2: Klick Dummy der Prototypen aus Figma.....	171
Anhang 2.1: Klick Dummy des Low-Fidelity-Prototype .....	171
Anhang 2.2: Klick Dummy des High-Fidelity-Prototype 1 .....	171
Anhang 2.3: Klick Dummy des High-Fidelity-Prototype 2 .....	171
Anhang 3: Dateien der Prototypen .....	171
Anhang 3.1: Datei des Low-Fidelity-Prototype.....	171
Anhang 3.2: Datei des High-Fidelity-Prototype 1 .....	171
Anhang 3.3: Datei des High-Fidelity-Prototype 2 .....	171
Anhang 3.4: Video / Bilder der KI-generierten Applikation.....	171
Anhang 4: Umfragen .....	171
Anhang 4.1: Umfrage Externes Benchmarking.....	171
Anhang 4.1.1: Umfrage Tee Kontor Kiel .....	171
Anhang 4.1.2: Umfrage The English Tearoom .....	172
Anhang 4.2: Umfrage-Low-Fidelity-Prototype .....	172
Anhang 4.3: Umfrage High-Fidelity-Prototype 1 .....	172
Anhang 4.4: Umfrage High-Fidelity-Prototype 2 .....	172
Anhang 4.5: Umfrage KI-generierte Applikation.....	172
Anhang 5: Mitschrift Erstellung der KI-generierten Applikation.....	172
Anhang 6: Einladung der Usability-Tests .....	172
Anhang 7: Mitschriften der Usability-Tests.....	172
Anhang 7.1: Mitschrift Low-Fidelity-Prototype.....	172
Anhang 7.2: Mitschrift High-Fidelity-Prototype 1 .....	172
Anhang 7.3: Mitschrift der KI-generierten Applikation.....	172
Anhang 7.4: Mitschrift der manuell gestalteten Applikation / High-Fidelity-Prototype 2 .....	172
Anhang 8: Methoden zur Erstellung der manuell gestalteten Applikation.....	173
Anhang 8.1: Externes Benchmarking: Bewertungen der Mitbewerber .....	173
Anhang 8.2: Informationsarchitektur .....	173
Anhang 8.3: Wireframes.....	173
Anhang 8.4: Styleguide .....	173

Anhang 8.5: Alle Methoden .....	173
Anhang 9: Excel Dateien .....	173
Anhang 9.1: Excel Datei Tee Kontor Kiel .....	173
Anhang 9.2: Excel Datei The English Tearoom .....	173
Anhang 9.3: Excel Datei Low-Fidelity-Prototype .....	173
Anhang 9.4: Excel Datei High-Fidelity-Prototype 1 .....	173
Anhang 9.5: Excel Datei der KI-generierte Applikation .....	174
Anhang 9.6: Excel Datei der manuellen gestalteten Applikation .....	174
Anhang 9.7: Excel Datei der Einzeldaten .....	174
Anhang 9.8: Excel Datei der Boxplot-Diagramme .....	174
Anhang 10: Berechnungen des Mann-Whitney-U-Tests .....	174
Anhang 11: Berechnungen des Medians und des Interquartilsabstand .....	174
Anhang 12: Testplan .....	174
Anhang 13. Exposé .....	174

## **Anhang 1: Fiktive Auftragsstellung**

Siehe separates Anhang Dokument

-> USB-Stick: Dateipfad: Dokumente\Anhang\_SeitzMarie

## **Anhang 2: Klick Dummy der Prototypen aus Figma**

### **Anhang 2.1: Klick Dummy des Low-Fidelity-Prototype**

<https://www.figma.com/proto/HiWVEqeHrvSVi7aNYoPrTq/LowFidelity?node-id=1-2&t=Khn678JbF-pRzjTVr-1>

### **Anhang 2.2: Klick Dummy des High-Fidelity-Prototype 1**

[https://www.figma.com/proto/NS6PfjyEXrvJ4koGNXVSzl/HighFidelity\\_1?node-id=1-2&t=haTdin7FCPXRIUAH-1](https://www.figma.com/proto/NS6PfjyEXrvJ4koGNXVSzl/HighFidelity_1?node-id=1-2&t=haTdin7FCPXRIUAH-1)

### **Anhang 2.3: Klick Dummy des High-Fidelity-Prototype 2**

[https://www.figma.com/proto/nIv9WirXWX8jqtZcThbchd/HighFidelity\\_2?node-id=1-2&t=vyEGLXJzlhwtvOgL-1](https://www.figma.com/proto/nIv9WirXWX8jqtZcThbchd/HighFidelity_2?node-id=1-2&t=vyEGLXJzlhwtvOgL-1)

## **Anhang 3: Dateien der Prototypen**

### **Anhang 3.1: Datei des Low-Fidelity-Prototype**

USB-Stick: Dateipfad: Prototypen\LowFidelityPrototype

### **Anhang 3.2: Datei des High-Fidelity-Prototype 1**

USB-Stick: Dateipfad: Prototypen\HighFidelityPrototype\_1

### **Anhang 3.3: Datei des High-Fidelity-Prototype 2**

USB-Stick: Dateipfad: Prototypen\HighFidelityPrototype\_2

### **Anhang 3.4: Video / Bilder der KI-generierten Applikation**

USB-Stick: Dateipfad: KIApplikation\KIApplikation\_Video

USB-Stick: Dateipfad: KIApplikation\KIApplikation\_Bilder

### **Anhang 3.5: Video der manuell gestalteten Applikation**

USB-Stick: Dateipfad: Prototypen\ManuelleApplikation\_Video

## **Anhang 4: Umfragen**

### **Anhang 4.1: Umfrage Externes Benchmarking**

#### **Anhang 4.1.1: Umfrage Tee Kontor Kiel**

<https://forms.gle/8PqMcTAsNV5J1g6V9>

## **Anhang 4.1.2: Umfrage The English Tearoom**

<https://forms.gle/xVsnGMgx56LYF4B6A>

## **Anhang 4.2: Umfrage-Low-Fidelity-Prototype**

<https://forms.gle/Mjp7Z6zUNPLtyVCY7>

## **Anhang 4.3: Umfrage High-Fidelity-Prototype 1**

<https://forms.gle/in9CpUd3g9S2u4WD6>

## **Anhang 4.4: Umfrage High-Fidelity-Prototype 2**

<https://forms.gle/h1Y6KpjoxisQw9Rk6>

## **Anhang 4.5: Umfrage KI-generierte Applikation**

<https://forms.gle/XPs1KDg149EGtxwV8>

## **Anhang 5: Mitschrift Erstellung der KI-generierten Applikation**

Siehe separates Anhang Dokument

-> USB-Stick: Dateipfad: Dokumente\Anhang\_SeitzMarie

## **Anhang 6: Einladung der Usability-Tests**

USB-Stick: Dateipfad: Dokumente\Einladung\_UsabilityTest

## **Anhang 7: Mitschriften der Usability-Tests**

### **Anhang 7.1: Mitschrift Low-Fidelity-Prototype**

Siehe separates Anhang Dokument

-> USB-Stick: Dateipfad: Dokumente\Anhang\_SeitzMarie

### **Anhang 7.2: Mitschrift High-Fidelity-Prototype 1**

Siehe separates Anhang Dokument

-> USB-Stick: Dateipfad: Dokumente\Anhang\_SeitzMarie

### **Anhang 7.3: Mitschrift der KI-generierten Applikation**

Siehe separates Anhang Dokument

-> USB-Stick: Dateipfad: Dokumente\Anhang\_SeitzMarie

### **Anhang 7.4: Mitschrift der manuell gestalteten Applikation / High-Fidelity-Prototype 2**

Siehe separates Anhang Dokument

-> USB-Stick: Dateipfad: Dokumente\Anhang\_SeitzMarie

## **Anhang 8: Methoden zur Erstellung der manuell gestalteten Applikation**

### **Anhang 8.1: Externes Benchmarking: Bewertungen der Mitbewerber**

Siehe separates Anhang Dokument

-> USB-Stick: Dateipfad: Dokumente\Anhang\_SeitzMarie

### **Anhang 8.2: Informationsarchitektur**

Siehe separates Anhang Dokument

-> USB-Stick: Dateipfad: Dokumente\Anhang\_SeitzMarie

### **Anhang 8.3: Wireframes**

Siehe separates Anhang Dokument

-> USB-Stick: Dateipfad: Dokumente\Anhang\_SeitzMarie

### **Anhang 8.4: Styleguide**

USB-Stick: Dateipfad: Methoden\Styleguide

-> USB-Stick: Dateipfad: Dokumente\Anhang\_SeitzMarie

### **Anhang 8.5: Alle Methoden**

USB-Stick: Dateipfad: Methoden\Alle\_Methoden

Miro Link:

[https://miro.com/welcomeonboard/ZTh1RVdoSkdWbUhDWm91ZTdNSFVzdWdnZmM5TVNFUmhQcGRwRzh0eFA5OVBOd1NRVyt3YnZPZDV1bnJCTW5oZkljWUJiMU5qWDVndWo5OUw5SWwzaTI2V1NSaXh5a0VnK3hkTWZpQ3IEaTY3VDh6bjgwMU9hK2ltbFgxcGd0OHFNakdSWkpBejJWRjJhRnhhb1UwcS9BPT0hdjE=?share\\_link\\_id=281533852598](https://miro.com/welcomeonboard/ZTh1RVdoSkdWbUhDWm91ZTdNSFVzdWdnZmM5TVNFUmhQcGRwRzh0eFA5OVBOd1NRVyt3YnZPZDV1bnJCTW5oZkljWUJiMU5qWDVndWo5OUw5SWwzaTI2V1NSaXh5a0VnK3hkTWZpQ3IEaTY3VDh6bjgwMU9hK2ltbFgxcGd0OHFNakdSWkpBejJWRjJhRnhhb1UwcS9BPT0hdjE=?share_link_id=281533852598)

## **Anhang 9: Excel Dateien**

### **Anhang 9.1: Excel Datei Tee Kontor Kiel**

USB-Stick: Dateipfad: Excel\Umfrage\_TeeKontorKiel

### **Anhang 9.2: Excel Datei The English Tearoom**

USB-Stick: Dateipfad: Excel\Umfrage\_TheEnglishTearoom

### **Anhang 9.3: Excel Datei Low-Fidelity-Prototype**

USB-Stick: Dateipfad: Excel\Umfrage\_LowFidelityPrototype

### **Anhang 9.4: Excel Datei High-Fidelity-Prototype 1**

USB-Stick: Dateipfad: Excel\Umfrage\_HighFidelityPrototype\_1

## **Anhang 9.5: Excel Datei der KI-generierte Applikation**

USB-Stick: Dateipfad: Excel\Umfrage\_KIApplikation

## **Anhang 9.6: Excel Datei der manuellen gestalteten Applikation**

USB-Stick: Dateipfad: Excel\Umfrage\_ManuelleApplikation

## **Anhang 9.7: Excel Datei der Einzeldaten**

USB-Stick: Dateipfad: Excel\Ergebnisse\_UsabilityTest\_Einzeldaten

## **Anhang 9.8: Excel Datei der Boxplot-Diagramme**

USB-Stick: Dateipfad: Excel\Boxplot\_Diagramme

## **Anhang 10: Berechnungen des Mann-Whitney-U-Tests**

USB-Stick: Dateipfad: Berechnungen\Mann-Whitney-U-Test

## **Anhang 11: Berechnungen des Medians und des Interquartilsabstand**

USB-Stick: Dateipfad: Berechnungen\Median-Interquartilsabstand

## **Anhang 12: Testplan**

USB-Stick: Dateipfad: Dokumente\Testplan\_SeitzMarie

## **Anhang 13. Exposé**

USB-Stick: Dateipfad: Dokumente\Expose\_SeitzMarie