

# **Hochschule Darmstadt**

– Fachbereich Informatik–

## **Vergleich derzeit aktueller Evaluationsmethoden zur Website Usability Optimierung**

Abschlussarbeit zur Erlangung des akademischen Grades  
Bachelor of Science (B.Sc.)

vorgelegt von

**Duy Van, Nguyen**

Matrikelnummer: 760265

Referent : Prof. Dr. Renz Kai  
Korreferent : Prof. Dr. Burda Daniel  
Bearbeitungszeit : 03.12.2021-03.03.2022

## ERKLÄRUNG

---

Ich versichere hiermit, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die im Literaturverzeichnis angegebenen Quellen benutzt habe.

Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder noch nicht veröffentlichten Quellen entnommen sind, sind als solche kenntlich gemacht.

Die Zeichnungen oder Abbildungen in dieser Arbeit sind von mir selbst erstellt worden oder mit einem entsprechenden Quellennachweis versehen.

Diese Arbeit ist in gleicher oder ähnlicher Form noch bei keiner anderen Prüfungsbehörde eingereicht worden.

*Darmstadt, 03.03.2022*



---

Duy Van, Nguyen

## ABSTRACT

# Einleitung

Vorgehen

Um eine Website benutzerfreundlich zu gestalten, reicht es nicht sich auf sein Bauchgefühl zu verlassen oder Annahmen zu treffen. Dafür gibt es Methoden, die systematisch herangehen um die Usability zu verbessern. Es stellt sich daher die Frage, welche Methoden derzeit für die Usability-Evaluation von Websites verwendet werden und wie sollte am besten ein geeignetes Verfahren aus dem breiten Angebot ausgewählt werden. Zunächst wird ein Überblick von dem aktuellen Forschungsstand geschaffen. Zwar gibt es drei verwandte Arbeiten zum Thema Usability-Evaluationsmethoden, jedoch hat jede Arbeit unterschiedliche Schwerpunkte. Daher wurde eine eigene Recherche angestellt, die 67 Studien aus den Jahren 2015 bis 2021 betrachtet und herausfindet, dass Parallelen zu den drei verwandten Arbeiten bestehen. So sind Heuristische Evaluation, Questionnaires (Fragebogen) und Automatisierte Tools, die derzeit meist verwendeten Methoden zur Evaluation von Website-Usability. Weitere Verfahren sind z. B.: Think-aloud, Eye-tracking, Interview, Usability-Testing, Cognitive Walkthrough und noch einige weitere. Im nächsten Schritt werden die Methoden aus der Recherche in drei große Kategorien klassifiziert: in expertenbasierte Methoden, in nutzerbasierte Methoden und in toolbasierte Methoden. Einige ausgewählte Methoden werden den Kategorien zugeordnet und genauer vorgestellt. Als nächstes werden die Methoden miteinander anhand von vorher bestimmten Kriterien verglichen. Es stellt sich heraus, dass die drei am häufigsten Verfahren aus der Recherche im Vergleich zu anderen Methoden mehr Vorteile bieten wie z. B. einfache Nutzung, schnelle Ergebnisse und geringer Ressourcenbedarf. Aufgrund der Vielzahl an Methoden wird ein Schema zur Unterstützung bei der Methodenauswahl erstellt. Im Anschluss erfolgt die Vorstellung einiger Tools, die gemeinsam mit den Methoden bei einer Evaluation der Usability unterstützend wirken können.

# Ergebnis

**Keywords:** Website-Usability, Usability-Evaluationsmethoden, expertenbasierte UEM, nutzerbasierte UEM, toolbasierte UEM, Heuristische Evaluation, Questionnaire, Automatisierte Tools, Eye-Tracking, Think-aloud, Usability-Testing, Interview, Cognitive Walkthrough

## INHALTSVERZEICHNIS

---

1	EINLEITUNG	1
2	VERWANDTE ARBEITEN	3
3	RESEARCH	5
3.1	Vorgehensweise . . . . .	5
3.2	Auswertung . . . . .	6
3.3	Zusammenfassung . . . . .	8
4	GRUNDLAGEN	9
4.1	Begriffserklärung . . . . .	9
4.1.1	Usability . . . . .	9
4.1.2	Usability-Metriken . . . . .	9
4.1.3	Usability-Probleme . . . . .	10
4.1.4	Usability-Evaluationsmethoden . . . . .	11
4.2	Klassifizierung der Usability-Evaluationsmethode . . . . .	11
4.3	Vorstellung Usability-Evaluationsmethoden . . . . .	12
4.3.1	Expertenbasierte Evaluationsmethoden . . . . .	13
4.3.2	Nutzerbasierte Evaluationsmethoden . . . . .	15
4.3.3	Toolbasierte Evaluationsmethode . . . . .	19
4.4	Zusammenfassung . . . . .	22
5	VERGLEICHE DER METHODEN	24
5.1	Bestimmung der Vergleichskriterien für UEM . . . . .	24
5.2	Vergleich der Usability-Evaluationsmethoden . . . . .	26
5.2.1	Vergleich der expertenbasierten UEM . . . . .	26
5.2.2	Vergleich der nutzerbasierten UEM . . . . .	29
5.2.3	Vergleich der toolbasierten UEM . . . . .	34
5.3	Zusammenfassung . . . . .	37
6	SCHEMA ZUR AUSWAHL VON EVALUATIONSMETHODEN	38
6.1	Bestimmung der Rahmenbedingungen . . . . .	38
6.2	Methodenauswahl . . . . .	38
7	TOOLS	41
7.0.1	Tools für nutzerbasierte UEM . . . . .	41
7.0.2	Tools für automatisierte Analyse . . . . .	41
8	SCHLUSS	46
8.0.1	Diskussion . . . . .	46
8.0.2	Zusammenfassung . . . . .	47
8.0.3	Ausblick . . . . .	48
	LITERATUR	49

## ABBILDUNGSVERZEICHNIS

---

Abbildung 1.1	Aufbau der Arbeit . . . . .	2
Abbildung 3.1	Suchbeispiel . . . . .	5
Abbildung 3.2	Darstellung der untersuchten Studien und Methoden .	6
Abbildung 3.3	Verteilung der Evaluationsformen [22] . . . . .	7
Abbildung 3.4	Nutzerbasierte Evaluationsmethoden [22] . . . . .	7
Abbildung 3.5	Meist verwendete Methoden in Software (2011-2014) [42] . . . . .	8
Abbildung 4.1	In den Studien behandelte Usability-Probleme [22] . .	11
Abbildung 4.2	Visualisierung der Daten als Scanpath und Heatmap .	18
Abbildung 4.3	Heatmaps . . . . .	21
Abbildung 4.4	beispielhaftes Usability Modell . . . . .	22
Abbildung 5.1	Empfohlene Anzahl der Tester (5-15) (Nielsen 2000) . .	31
Abbildung 5.2	Klick Analyse Methodik [31] . . . . .	35
Abbildung 5.3	technischer Hinweis . . . . .	36
Abbildung 6.1	Schema zur Auswahl der Methode . . . . .	40
Abbildung 7.1	Resultat der Website <a href="https://fbi.h-da.de/">https://fbi.h-da.de/</a> durch Qua- lidator . . . . .	42
Abbildung 7.2	Resultat der Website <a href="https://fbi.h-da.de/">https://fbi.h-da.de/</a> durch GT- mextric . . . . .	43
Abbildung 7.3	Resultat der Website <a href="https://fbi.h-da.de/">https://fbi.h-da.de/</a> durch Nibbler	44
Abbildung 7.4	Resultat der Website <a href="https://fbi.h-da.de/">https://fbi.h-da.de/</a> durch Po- werMapper . . . . .	45

## ABKÜRZUNGSVERZEICHNIS

---

HE	Heuristische Evaluation
CW	Cognitive Walkthrough
ET	Eye-Tracking
UEM	Usability-Evaluationsmethode
UEMs	Usability-Evaluationsmethoden
HCI	human-computer interaction
ISO	International Standards Organization
UXPA	Usability Professional Association
SDLC	Software development lifecycle
SUS	System Usability Scale
QUIS	Questionnaire for User Interface Satisfaction
WAMMI	Website Analysis and measurement inventory
CSUQ	Computer System Usability Questionnaire
NAU	Nielsen 's Attributes of Usability
HFRG	Human Factors Research Group
AOI	Areas of Interest

## EINLEITUNG

---

Wir leben in einer vernetzten Welt. Das Web bietet vielfältige Online-Dienste um das tägliche Leben bequemer zu machen. Die Nutzung von Websites nimmt enorm zu für alle Arten von Menschen. Die erste Website wurde vor ca. 30 Jahren von Tim Berners Lee veröffentlicht. Zwischen dem Jahr 2000 und 2010 ist ein kontinuierliches Wachstum der Webseiten festzustellen von ca. 20 Millionen auf rund 240 Millionen Webseiten. 2010 setzte ein exponentielles Wachstum ein. Im Jahr 2021 gibt es bereits 1,88 Milliarden Webseiten [10]. Unternehmen werben über ihre Website für ihre Produkte und Online-Dienste. Daher ist es essentiell, dass bei der Gestaltung der Website die Bedürfnisse der Nutzer im Mittelpunkt stehen. Ist die Nutzung der Website nicht intuitiv, kann das Frustrations- sowie Unzufriedenheit bei den Nutzern hervorrufen. Das kann zur Folge haben, dass Nutzer nach kurzer Zeit die Seite verlassen und nach besseren Alternativen suchen. Aus diesem Grund spielt eine ausgereifte Gestaltung einer Website für Unternehmen, die ihre Dienste über das Netz anbieten, eine wichtige Rolle. Eine zentrale Rolle spielt auch die Usability einer Website, denn die Benutzerfreundlichkeit ist ein bedeutender Faktor für den Erfolg einer Website. Damit eine Website sich von anderen Websites abheben kann, bedarf es nicht nur einer einmaligen Entwicklung. Websites müssen regelmäßig gepflegt und kontinuierlich verbessert werden. Um dies zu erreichen sind verschiedene Techniken nötig, um Schwachstellen im Design zu identifizieren und die Websites damit nutzerfreundlicher zu gestalten.

Es handelt sich bei dieser Arbeit um eine Vergleichsarbeit. In der Arbeit wird der Frage nachgegangen, *welche Evaluationsmethoden derzeit verwendet werden, um die Usability von Websites zu verbessern*. Es werden dabei drei Subfragen formuliert um das Problem systematisch Schritt für Schritt zu bearbeiten.

- 1) *Welche Evaluationsmethoden werden derzeit in Studien eingesetzt um die Usability von Webseiten zu evaluieren?*
- 2) *Worin unterscheiden sich die ausgewählten Methoden untereinander?*
- 3) *Wie ist die Vorgehensweise um eine Methode aus dem breiten Angebot auszuwählen?*

Um die erste Subfrage zu beantworten wird zunächst nach verwandten Arbeiten recherchiert, um sich einen Überblick über die aktuellen Stand der Literatur zu verschaffen. Danach wird eine Recherche nach aktuellen Studien aus den Jahren 2015 bis 2021 mit Bezug auf Website-Usability Evaluation angestellt um derzeit angewandte Methoden zu identifizieren.

Dann wird ein Ranking der am meisten verwendeten Verfahren erstellt und diese einzeln vorgestellt. Anschließend wird der zweiten Subfrage nachgegangen, indem ein detaillierter Vergleich dieser Methoden anhand von vorher erstellten Vergleichskriterien gemacht wird.

Die dritte Frage wird beantwortet, indem ein Schema mit Hilfe des Vergleichs als weitere Hilfestellung zur Methodenauswahl erstellt wird. Zum Schluss kommt die Vorstellung einiger Tools, als Unterstützung der Methoden beim Evaluieren. Folgend ist der Aufbau der Arbeit visuell dargestellt.

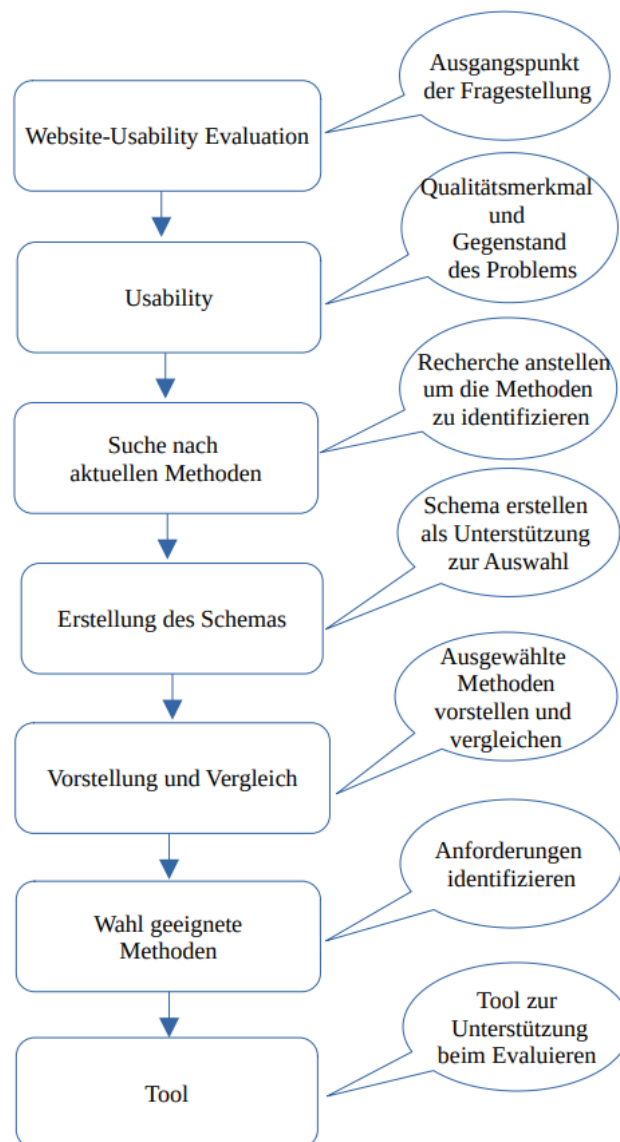


Abbildung 1.1: Aufbau der Arbeit



## VERWANDTE ARBEITEN

---

Diese Arbeit trägt den Titel: **Vergleich derzeit aktueller Evaluationsmethoden zur Website-Usability Optimierung**. Das bedeutet, es handelt sich bei dieser Arbeit um einen *Vergleich*, von *aktuellen Evaluationsmethoden* mit dem Ziel die *Usability* von Websites mit geeigneten Methoden zu verbessern.

Es ist in erster Linie wichtig, dass die Literatur rund um diese Thematik auf dem aktuellen Stand sind. Zu den Literaturen zählen Studien, Papers, wissenschaftliche Artikel, wissenschaftliche Bücher etc... Es gibt bereits ein paar Quellen, die sich mit den aktuellen Usability-Methoden auseinandersetzen haben wie z. B. :

- "Current Trends in Usability Evaluation Methods: A Systematic Review" [42] aus dem Jahr 2014
- "Research Trends in Web Site Usability: A Systematic Review" [22] aus dem Jahr 2016
- "The State of Art in Website Usability Evaluation Method " [85] aus dem Jahr 2017

Doch was bedeutet genau "aktueller Stand". Ab welchem Jahr gilt die Literatur als aktuell? Diese Frage lässt sich nicht einfach pauschal beantworten. Auf den ersten Blick sind die Literaturen im Jahr 2016 und 2017 aktueller die Literatur aus dem Jahr 2014. Doch es hängt davon ab, in welchem Zeitraum die Literatur die Evaluationsmethoden untersucht hat.

Die Literatur "The State of Art in Website Usability Evaluation Method" [85] (2017) untersuchte insgesamt 90 Research-Papers aus den letzten 20 Jahren. 20 Jahre sind ein langer Zeitraum und daher stellt diese Literatur auch eine Vielzahl an Evaluationsmethoden vor. Der Schwerpunkt dieses Papers liegt auf dem Verständnis verschiedener verwendeter Usability-Evaluationsmethoden (UEMs), die zur Evaluation von Websites benutzt werden können, daher ist es aus der Literatur nicht möglich zu entnehmen, welche Methoden beispielsweise in den letzten fünf oder zehn Jahren verwendet wurden. Sie bietet jedoch einen großen Überblick darüber, welche UEMs es gibt und was unter jeder Methode zu verstehen ist.

Die Literatur "Current Trends in Usability Evaluation Methods: A Systematic Review" [42] (2014) analysierte systematisch insgesamt 274 Usability-Studien in den drei Jahren zwischen 2011 und 2015 und veröffentlichte eine Tabelle, wie hoch der Anteil der jeweiligen Methode ist. Doch die Statistik beinhaltet nicht nur UEMs für Software-Produkte aus dem Bereich Websites auf, sondern auch aus dem anderen Bereichen wie z. B. Software zur Behandlung von Krankheiten, Software für Ärzte oder Software zur Durchführung einer Operation...[42]. UEMs für Websites ist nur ein kleiner Teil in

der Statistik. Die Arbeit hat aber einen systematischen Ansatz zur Ermittlung von aktuellen UEMs. Daher ist es nachvollziehbar, wie die Statistik zu Stande gekommen ist.

Die Studie "Research Trends in Web Site Usability: A Systematic Review" [22] (2016) wählte auch einen systematischen Ansatz und untersuchte insgesamt 199 Studien von 2005 bis 2014 und bezog sich nur auf UEMs für Websites. Die Auswertung dieser Recherche ist umfangreicher als die anderen beiden oben genannten Arbeiten. Die Recherche stellte mehrere wertvolle Ergebnisse vor wie z. B. wie viele Studien aus welchem Jahr untersucht worden sind, welche Gruppen von Evaluationsmethoden am meisten verwendet wurden und welche Methode am häufigsten zum Einsatz kam.

Keine der drei Studien erfüllt die vollständige Zielsetzung der Arbeit, wenn es darum geht, aktuelle Methoden im Bereich Website-Usability zu finden. Mal untersuchte eine Studie Evaluationsmethoden allgemein für Softwareprodukte, mal liegt der untersuchte Zeitraum 2005 bis 2014 sehr weit zurück oder gar zwei Jahrzehnten, mal liegt der Schwerpunkt auf dem Verständnis der verwendeten Methoden. Deswegen wurde dafür eine eigene Studie erstellt und diese wird im nächsten Kapitel vorgestellt.

## RESEARCH

Bei diesem Kapitel handelt es sich um eine Studie über die Häufigkeit der Evaluationsmethoden von Websites in den letzten sechs Jahren. Dafür wurden 67 Studien untersucht, die sich mit der evaluation der Usability von Websites befasst haben. Die Ergebnisse dienen als Grundlage für die Vorstellung der am meisten verwendeten Methoden sowie für den Vergleich der aktuellen Usability-Evaluationsmethoden.

## 3.1 VORGEHENSWEISE

Für die Untersuchung wird auch der systematischen Ansatz, wie bei der Studie [42], gewählt. Das Ziel ist es, den aktuellen Trend der bei der Verwendung von Usability-Evaluationsmethoden (UEMs) herauszufinden. Damit soll die Studie die folgende Frage beantworten:

*Welche Methoden wurden in den letzten sechs Jahren (2015-2021) zur Usability-Evaluation von Websites am häufigsten verwendet?*

Der Ansatz für die Untersuchung ist an die Herangehensweise der Studie [42] angelehnt. Für die systematische Recherche wurde die Suchmaschine für Literaturen *Semantic Scholar* [89] verwendet und als Suchbegriff wurde die Wortgruppe *usability evaluation website study* gewählt. Die Suchergebnisse wurden nach Relevanz (Sort by Relevance) sortiert und das Erscheinungsjahr (Date Range) beginning mit 2015 iterativ bis 2021 fortgeführt. Damit soll sicher gestellt sein, dass nur aktuelle (Date Range) und relevante (Relevance) Studien in die Analyse aufgenommen werden. Die Abbildung 3.1 (Sucheispiel) zeigt beispielhaft wie die Suche durchgeführt wurde.

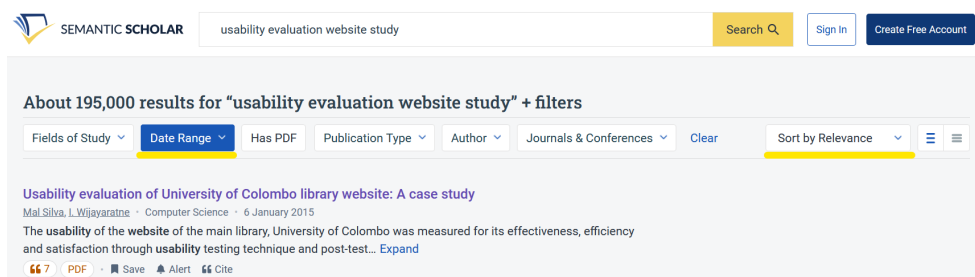
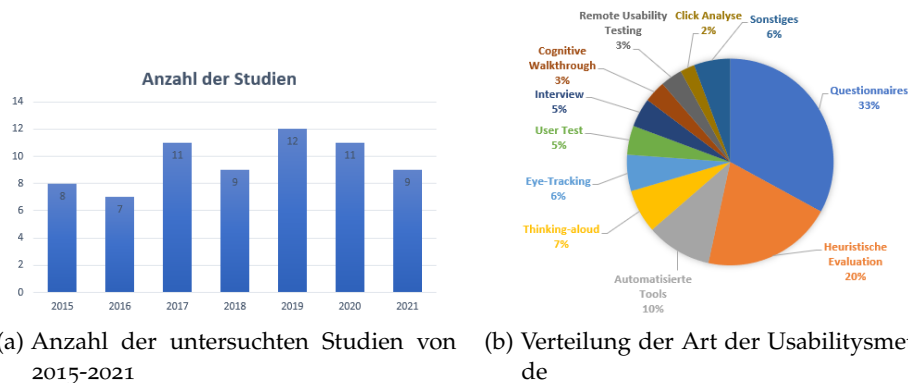


Abbildung 3.1: Sucheispiel

## 3.2 AUSWERTUNG

Insgesamt sind 67 Studien mit dem Schwerpunkt auf Usability Evaluierungsmethoden nur für Websites, untersucht worden. Alle untersuchten Studien liegen in PDF-Form zum Nachvollziehen der Auswertung vor. Die Anzahl der Studien aus den entsprechenden Jahren wurde, soweit es möglichst war, gleich gehalten. Im Durchschnitt wurden zehn Studien aus jedem Jahr für die Analyse gewählt. Die Abbildung 3.2 (a) (Anzahl der untersuchten Studien von 2015-2021) zeigt die genaue Verteilung.



(a) Anzahl der untersuchten Studien von 2015-2021 (b) Verteilung der Art der Usabilitysmethode

Abbildung 3.2: Darstellung der untersuchten Studien und Methoden

Es wurden in den 67 Studien insgesamt 15 unterschiedliche Evaluationsmethoden und 89 Einsätze festgestellt. Die 89 Einsätze (Abbildung 3.2 b) kommen zustande, weil in einigen Studien mehr als eine Evaluationsmethode verwendet wurden. Folgend sind die Methoden mit der Anzahl der Einsätze absteigend sortiert, aufgelistet.

- Questionnaires (Fragebögen): 29
- Heuristische Evaluation: 18
- Automatisierte Tools: 9
- Think-aloud: 6
- Eye-Tracking: 5
- Interview: 4
- Usability Testing: 4
- Cognitive Walkthrough: 3
- Remote Usability Testing: 3
- Klick-Analyse: 2
- Sonstige Methoden: 6

Fragebögen, Heuristische Evaluation und Automatisierte Tools sind die am häufigsten verwendeten Methoden aus den letzten sieben Jahren.

Eine Studie von 2016 [22] hat 199 Studien aus dem Jahre 2005 bis 2014 analysiert und fand heraus, dass nutzerbasierte Evaluationsmethoden (Abbildung 3.) in Usability Studien mit Abstand am meisten verwendet werden. Es sind 138 von 199 Studien. Das entspricht einem Anteil von ca. 70% aller Studien. Die folgende Grafik zeigt die Verteilung der Evaluationsformen. Aus der eigenen Analyse heraus, beträgt dieser Anteil ca. 63% (56 von 88 Einsätze waren nutzerbasierte Evaluationsmethoden). Es besteht somit eine Parallele zu dieser Studie.

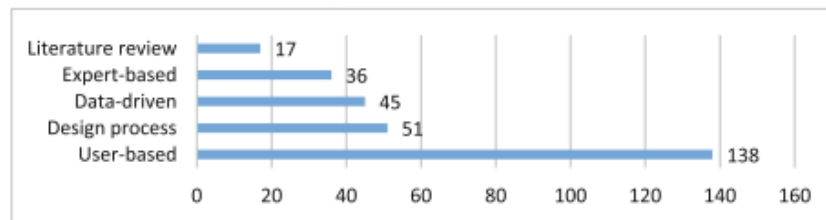


Abbildung 3.3: Verteilung der Evaluationsformen [22]

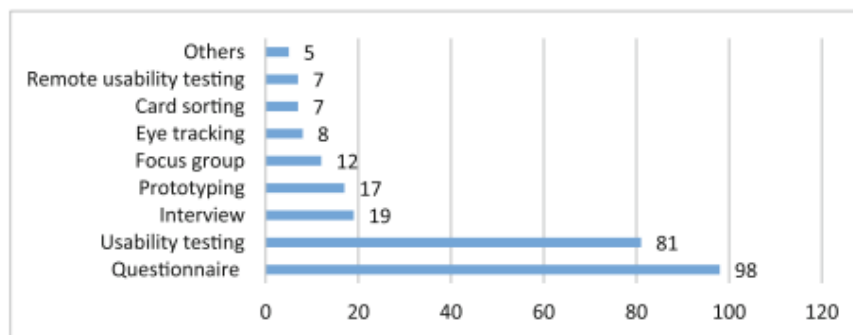


Abbildung 3.4: Nutzerbasierte Evaluationsmethoden [22]

Des Weiteren ist zu erkennen, dass die Methode Questionnaire (Abbildung 3.4 und Abbildung 3.5) im Bereich *nutzerbasierte Methoden* in beiden Studien die am meisten verwendete war. Auch ein Vergleich der Ergebnisse aus der Analyse mit der folgenden Grafik erlaubt es Gemeinsamkeiten festzustellen. So liegen Heuristische Evaluation und Questionnaire auch vorne bei den meist verwendeten Evaluationsmethoden, mit dem Unterschied, dass die Grafik sich (Abbildung 3.5) allgemein auf Software bezieht und die Analyse ausschließlich auf Websites. Des weiteren fällt auf, dass automatisierte Tools als Methode zur Usability Evaluation von Websites populär sind. Das liegt vermutlich daran, dass diese Methode einfach auf Websites anzuwenden ist.

Usability Method	Amount of evaluations	Percentage (%)
Usability Testing	83	25.70
Questionnaire	74	22.91
Heuristic Evaluation	54	16.72
Thinking Aloud	31	9.60
Interview	18	5.57
Prototype Evaluation	15	4.64
Automated Evaluation with Tool	12	3.72
Usability Metrics	10	3.10
Focus Group	7	2.17
Cognitive Walkthrough	6	1.86
Field Observation	5	1.55
Others (Co-discovery, GOMs Analysis, Guideline Review, Cognitive Task Analysis, Personas)	8	2.48
<b>Total</b>	<b>323</b>	<b>100.00</b>

Abbildung 3.5: Meist verwendete Methoden in Software (2011-2014) [42]

### 3.3 ZUSAMMENFASSUNG

Die Recherche wurde durchgeführt mit dem Ziel die Methoden zu identifizieren, die derzeit am meisten zur Evaluation der Usability von Websites verwendet werden. Für die Recherche wurde ein systematischer Ansatz gewählt. Dabei wurden 67 Studien aus den Jahren 2015 bis 2021 analysiert und die drei meist benutzten Methoden sind: Questionnaire, Heuristische Evaluation und Automatisierte Tools. Es fällt auf, dass Questionnaire und Heuristische Evaluation wie bei den anderen zwei Arbeiten [42], [22] die Tabelle führen. Die Verwendung von Automatisierten Tools ist in den letzten Jahren gestiegen.

## GRUNDLAGEN

---

### 4.1 BEGRIFFSERKLÄRUNG

Dieses Teilkapitel definiert einige zentrale Begriffe rund um das Thema Evaluation und Usability, damit soll ein Grundverständnis für die nächsten Kapitel geschaffen werden.

#### 4.1.1 Usability

Der Begriff *Usability* hat seinen Ursprung nicht in der Informatik. Er kann in verschiedenen Kontexten für jede Art von Produkt mit Menscheninteraktionen verwendet werden. Laut International Standards Organization (ISO) 9241 in Part 11: Guidance on Usability ist die Usability zu verstehen als Ausmaß, in dem ein System, Produkt oder eine Dienstleistung von bestimmten Benutzern verwendet werden kann, um bestimmte Ziele *effektiv*, *effizient* und *zufriedenstellend* in einem bestimmten *Nutzungskontext* zu erreichen[23]. Der Nutzungskontext ist laut ISO 9241-11 definiert als „Die Benutzer, Arbeitsaufgaben, Ausrüstung (Hardware, Software und Materialien) sowie physische und soziale Umgebung, in der das Produkt genutzt wird“. Für den Kontext dieser Arbeit kann die Usability als die Benutzbarkeit, die den Nutzern ermöglicht, deren Ziele bei der Bedienung einer Webseite zu erreichen, verstanden werden.

Es ist hierbei wichtig, die Software zusammen mit den Nutzern und den Zielen in einem sinnvollen Nutzungskontext steht. Das Augenmerk liegt hier auf dem Nutzungskontext, weil wenn z. B. sich eine perfekte Software mit den perfekten Zielen an die falsche Zielgruppe richtet, ist die Software somit nicht optimal.

#### 4.1.2 Usability-Metriken

*Usability-Metriken* dienen zur Messung von Userinteraktion mit einem Produkt. ISO-9126 und Nielsen gehen von folgenden Metriken aus [71], [63].

**EFFEKTIVITÄT (EFFECTIVENESS)** Die Effektivität bezieht sich auf die Ziele, die erreicht werden können, bei der Verwendung des Produkts mit der Genauigkeit und Vollständigkeit.

**EFFIZIENZ (EFFICIENCY)** Die Effizienz setzt den Grad der Effektivität zur Menge der eingesetzten Ressourcen fest. Im Allgemeinen kann die Effizienz durch die Zeit, welche zum Erfüllen der Aufgabe, benötigt wird, gemessen werden.

**ZUFRIEDENHEIT (SATISFACTION)** Die Zufriedenheit beschreibt die subjektive Reaktion eines Benutzers bei der Verwendung eines Produkts. Benutzer Zufriedenheit kann mit der Motivation, ein Produkt zu benutzen, im Zusammenhang stehen. Das System sollte angenehm zu bedienen sein.

**ERLERNBARKEIT (LEARNABILITY)** Das System sollte einfach zu erlernen sein, damit der Benutzer schnell mit seiner Arbeit beginnen kann.

**EINPRÄGSAMKEIT (MEMORABILITY)** Das System sollte einfach zu verwenden sein, sodass er selbst nach einer Zeit immer noch in der Lage ist, die Anwendung zu bedienen.

**FEHLER (ERRORS)** Das System sollte eine geringe Fehlerrate haben, sodass der Benutzer auf wenige Fehler stößt.

#### 4.1.3 Usability-Probleme

Um die Usability eines Produkts verbessern zu können, müssen *Usability-Probleme* bekannt sein. Wie so oft, gibt es keine eindeutige Definition was dieser Begriff bedeutet. Jedoch gibt es einige Studien, bei denen die Definitionen größtenteils übereinstimmen. Die Autoren der Fachzeitschrift "Behaviour & Information Technology" definiert Usability-Probleme wie folgt:

*"A usability problem is an aspect of the system and/or a demand on the user which makes it unpleasant, inefficient, onerous or impossible for the user to achieve their goals in typical usage situations"*[16].

Eine weitere Definition aus dem Jahr 2016 hat eine ähnliche Auffassung:

*"A usability problem is a set of negative phenomena, such as user's inability to reach his/her goal, inefficient interaction and/or user's dissatisfaction, caused by a combination of user interface design factors and factors of usage context."*[38].

Für einen Artikel [50] verwenden die Autoren die Definition von Usability-Problems als *"anything that interfered with a user's ability to efficiently and effectively complete tasks"*.

Diese drei Definitionen haben viele Gemeinsamkeiten. Beispielsweise beinhalten die Definitionen Begriffe wie Nutzer, Ziel, Interaktion, und Ineffizienz. In allen drei Definitionen steht der Nutzer im Mittelpunkt. Zusammenfassend können Usability-Probleme so verstanden werden, wenn Nutzer beim Interagieren mit einem Produkt seine Ziele nicht effizient erreichen können, dann ist das ein Usability Problem. Auf den Website Usability Kontext übertragen, bedeutet das, dass Nutzer beim Bedienen einer Webseite Schwierigkeiten haben, ihre Ziele zu erreichen und sie dabei unzufrieden sind.

Konkrete Usability-Probleme können z. B. sein: die Ästhetik, die Suche, UI-Design, die Navigation usw... Die folgende Grafik zeigt, die Usability Problems, die am meisten in Studien zwischen 2005-2015 behandelt wurden[22].



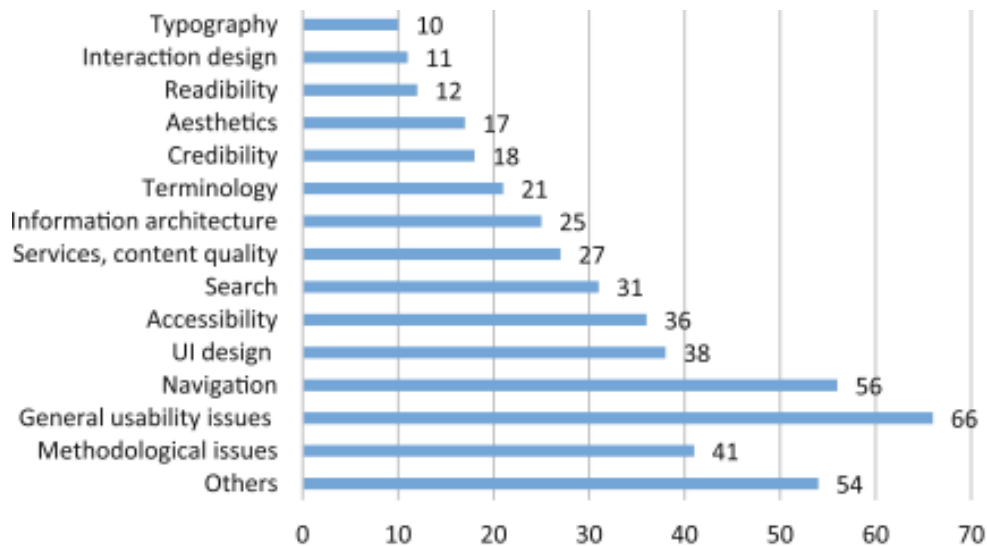


Abbildung 4.1: In den Studien behandelte Usability-Probleme [22]

#### 4.1.4 Usability-Evaluationsmethoden

Die Usability (dt. Benutzerbarkeit, Benutzerfreundlichkeit) ist ein Qualitätsmerkmal eines Produkts. Das führt unmittelbar dazu, dass Interessen entstehen, Usability-Evaluationsmethode (UEM) zu entwickeln, die dabei helfen sollen, die Usability zu evaluieren und Usability-Probleme zu identifizieren.

*Evaluation* ist an den englischen Begriff *evaluate* in der Bedeutung von *beurteilen, begutachten* angelehnt [19]. Der Begriff Evaluation wird von der *Joint Committee On Standards For Educational Evaluation* definiert als “die systematische Untersuchung der Verwendbarkeit oder Güte eines Gegenstandes” [19]. Der Definition nach [19] können die Gegenstände ein Programm, ein Projekt oder Materialien sein. Auf diese Arbeit übertragen, ist die Evaluation die systematische Untersuchung der Usability von Websites.

Zusammen mit den Begriffen Usability und Methode versteht sich der Begriff *Usability-Evaluationsmethode* als jede Methode, die verwendet wird, um die Usability von Websites systematisch zu untersuchen.

## 4.2 KLASSIFIZIERUNG DER USABILITY-EVALUATIONSMETHODE

Über die Jahre sind viele Methoden entstanden. Mittlerweile steht eine breites Spektrum an Evaluationsmethoden zur Verfügung. Zu diesem Spektrum gehören beispielsweise: Usability Testing, Think-aloud, Fokusgruppe, Web Analytics, heuristische Evaluation, Eye-Tracking, Fragebögen, Card Sorting, Interviews...[85], [22]. Es gibt noch weitaus viel mehr Methoden. Um sich über diese breite Masse an Usability-Evaluationsmethoden, die über die Jahre entstanden sind, einen besseren Überblick verschaffen zu können, ist es von Vorteil eine geeignete Gruppierung zu erstellen. Die Klassifizierung dient als Ausgangspunkt zur Auswahl geeigneter Methoden für bestimmte Anwendungsfelder.

Es gibt schon einige Autoren, welche sich mit der Klassifizierung auseinander gesetzt hat. Die Studie [85] hat bei drei Autoren drei Varianten der Klassifizierung aufgezeigt z. B. Hasan (2009) unterteilt die UEM in evaluatorenbasierte, nutzerbasierte und toolbasierte UEM; Ivory (2001) und Fernanderz (2011) klassifiziert in fünf Kategorien. Eine Studie, welche 199 Studien von Website UEM zwischen 2005 und 2014 untersucht hat, schlägt eine Klassifizierung in fünf Hauptkategorien vor: nutzerbasierte UEM, expertenbasierte UEM, Data-driven, Design process, Literatur review.

Doch nach welchem Merkmal sollen die Evaluationsmethoden eingruppiert werden. Eine Arbeit [28] hat mehrere Merkmale auf die Verfahren bezogen herausgearbeitet, die dabei helfen, die Verfahren systematisch einzuordnen. Der Arbeit nach gibt es zwei große Klassifizierungen: *dimensionale Klassifizierung* und *kategoriale Klassifizierung*. In der dimensionalen Klassifizierung können die Evaluationsmethoden nach: *Erscheinungsjahr*, *Erscheinungsstandort*, *Erfinder oder Herausgeber* eingeordnet werden. In der kategoriale Klassifizierung ist die Einsortierung nach: *Untersuchungszweck* (formative und summativ Evaluation), *Untersuchungskriterien* (Usability DIN ISO 9241-11 1998), *Analyseform* (quantitative und qualitative Methoden) oder *Gutachter/Evaluator* möglich. Für diese Arbeit wird die Klassifizierung nach Gutachter/Evaluator gewählt, damit die Nutzer der Software im Mittelpunkt stehen und bei der Klassifizierung mitberücksichtigt werden. Demnach können die Evaluationsmethoden in drei Kategorien [2], [43] eingeteilt werden: expertenbasierte UEM, nutzerbasierte UEM und toolbasierte UEM.

**EXPERTENBASIERTE EVALUATIONSMETHODEN** : Um die Usability eines Systems zu messen und zu bewerten, nehmen an dieser Kategorie der Usability-Evaluierung eine Reihe von Usability-Experten teil [84]. Laut der Aussage dieser Studie [84] gelten Heuristische Evaluation und Cognitive Walkthrough als die am weitesten verbreitete Methode in dieser Kategorie. Die Aussage kann durch die eigene Untersuchung bestätigt werden (Abbildung 3.2).

**NUTZERBASIERTE EVALUATIONSMETHODEN** : Websites oder allgemein interaktive Software-Produkte sind an die Endnutzer gerichtet. Aus dem Grund spielen Endnutzer eine wichtige Rolle und sie müssen bei der Evaluation miteinbezogen werden.

**TOOLBASIERTE EVALUATIONSMETHODEN** : Diese Art von Methoden wird verwendet, um Usability Probleme automatisch zu identifizieren mit unterschiedlichen Tools.

#### 4.3 VORSTELLUNG USABILITY-EVALUATIONSMETHODEN

Die am meisten verwendeten Evaluationsmethoden aus der Recherche werden nun in die entsprechenden Klassen einsortiert und die einzelnen Verfahren werden näher erklärt um ein besseres Verständnis zu schaffen. Die UEMs sind: Questionnaire, Heuristische Evaluation, Automatisierte Tools, Think-

aloud, Eye-Tracking, Interview, Usability Testing, Cognitive Walkthrough, Remote Usability Testing und Klick-Analyse.

#### 4.3.1 Expertenbasierte Evaluationsmethoden

##### 4.3.1.1 Heuristische Evaluation

Laut der Usability Professional Association (UXPA) Umfrage 2010 ist Heuristische Evaluation die zweit meist benutzte Methode bei Organisationen weltweit [96]. Der Prozess der Identifizierung potenzieller Usability-Probleme durch den Vergleich eines Benutzeroberflächendesigns mit den etablierten Usability-Design-Richtlinien werden als Expertenevaluation oder auch als *Heuristische Evaluation* bezeichnet[18]. Eine Heuristische Evaluation (HE) ist eine systematische Inspektion einer Benutzeroberfläche durch eine kleine Gruppe von Evaluatoren [88]. Bei der HE prüfen die Experten (Evaluatoren), inwiefern das Produkt die gewünschten Heuristiken (Prinzipien) aufweist. Molich und Nielsen haben im Jahr 1990 zehn Usability Heuristiken veröffentlicht [8] und im Jahr 2020 [41] aktualisiert. Die von Nielsen entwickelten Heuristiken sind eine gute Grundlage für die Usability-Evaluation mit dieser Methode. Mit der Zeit werden Anwendungen immer komplexer und es gibt ganz unterschiedliche Anwendungen, daher müssen die Heuristiken von Fall zu Fall angepasst oder gar selbst erweitert werden. So verwendet die Studie [12] insgesamt 15 Heuristiken. Auch wenn diese Methode bereits im Jahr 1990 entstanden ist, wird sie in der heutigen Zeit immer noch zur Webseite Evaluation eingesetzt (Abbildungen 3.2 und 3.5) und ist eine der am verbreitetsten Usability Evaluationsmethode [95]. Laut Nilsen und Molich wird Heuristische Evaluation so umgesetzt, indem die Meinung durch Evaluatoren abgegeben wird, was an der Schnittstelle gut und schlecht ist [77], unter Einsatz von Usability Prinzipien [88]. Folgend sind zehn von Nielsen entwickelte Heuristiken: [100], [41], [64].

1. *Sichtbarkeit des Systemstatus*: Das Design sollte die Benutzer immer über das Geschehen auf dem Laufenden halten.
2. *Übereinstimmung zwischen System und realer Welt*: Das Design sollte die Sprache der Nutzer sprechen. Wörter, Ausdrücke und Konzepte, die den Nutzern vertraut sind, sollten verwendet werden. Die Informationen sollten in einer natürlichen und logischen Reihenfolge erscheinen.
3. *Benutzerkontrolle und Benutzerfreiheit*: Benutzer führen oft versehentlich Aktionen aus. Ihnen sollte die Möglichkeit gegeben werden, die ungewollte Aktion schnellstmöglich rückgängig machen zu können.
4. *Konsistenz und Standards*: Das Design sollte den Plattform- und Branchenkonventionen entsprechen.
5. *Fehlervermeidung*: Gute Fehlermeldungen sind wichtig, aber besser ist es, die Designs so zu entwerfen, dass Probleme überhaupt erst nicht entstehen.

6. *Erkennen statt Erinnern*: Nutzer sollten sich nicht Informationen von einem Teil der Schnittstelle zu einem anderen merken müssen. Elemente, Aktionen und Optionen sollten leicht zu erkennen sein.
7. *Flexibilität und Effizienz der Nutzung*: Shortcuts, die für unerfahrende Nutzer verborgen sind, können die Interaktion für erfahrene Nutzer beschleunigen, sodass das Design sowohl für unerfahrene als auch für erfahrene Nutzer geeignet ist.
8. *Ästhetisches und minimalistisches Design*: Schnittstellen sollten keine irrelevanten Informationen enthalten. Diese können Nutzer ablenken, daher sollte der Fokus beim Design der Oberfläche auf dem Wesentlichen liegen.
9. *Nutzerfreundliche Fehlermeldungen*: Fehlermeldungen sollten in Klartext formuliert sein und das Problem beschreiben.
10. *Support und Dokumentation*: Es ist am besten, wenn das System keiner zusätzlichen Erklärung bedarf. Jedoch kann es in vielen Fällen nützlich sein, eine Dokumentation bereitzustellen, die den Nutzern bei Fragen hilft.

Diese Prinzipien sind zwar allgemein definiert, doch sie können als Richtlinien dienen, gegen die nicht verstoßen werden sollte.

Eine ganz wichtige Rolle nehmen die Experten bei der HE ein. Sie zeichnen sich durch fundierte Kenntnisse auf dem Gebiet human-computer interaction (HCI) aus [88]. Die Auswahl der Experten für die Evaluation ist besonders wichtig, weil erfahrene Experten die meisten Usability-Probleme erkennen können. Nielsen fand heraus, dass Experten mit besonderer Expertise in Systemoberfläche und Usability-Attribute erkennen im Vergleich zu Experten mit beispielsweise nur Expertise in Usability-Attributen deutlich mehr Usability-Probleme. [88]. Bereits mit fünf Experten können ca. 75% der Usability-Probleme aufgedeckt werden, bei 15 Experten sind es ca. 90% aber die optimale Anzahl der Experten liegt zwischen drei und fünf [88].

Wenn die Usability-Probleme identifiziert wurden, sollen diese nach unterschiedlichem Schweregrad einsortiert werden, z. B. mithilfe einer 5-Punkte-Skala [88]. Nielsen schlägt die folgende 5-Punkte-Skala vor, um die Usability-Probleme zu bewerten.

- 0: es ist kein Usability-Problem
- 1: *kosmetisches Problem*, beheben wenn es Zeit dafür gibt
- 2: *minor Usability-Problem*, hat niedrige Priorität
- 3: *major Usability-Problem*, hat hohe Priorität
- 4: *katastrophe Usability*, muss unbedingt behoben werden

Damit können die Usability-Probleme priorisiert werden. Je kleiner die Zahl desto weniger Priorität hat das jeweilige Problem.

#### 4.3.1.2 Cognitive Walkthrough

*Cognitive Walkthrough* ist eine der am weitesten verbreiteten expertenbasierten Methoden, die auf die einfache Erlernbarkeit eines System fokussiert ist [51]. Es bietet eine Möglichkeit das User-Interface und ihre Lernfähigkeit durch Experten, die sich in die Rolle der Nutzer versetzen, zu testen [59]. Cognitive Walkthrough (CW) wurde in den frühen 1990er Jahren entwickelt [91]. Das Verfahren ist eine aufgabenbezogene Evaluationsmethode, bei der die Experten eine Reihe von Aufgabenszenarien durcharbeiten [37]. Das Identifizieren von Usability-Problemen wird durch die Simulation von Aufgaben unterschiedlichen Schwierigkeitsgrades realisiert [91]. CW besteht aus zwei Phasen: die Vorbereitungsphase und die Ausführungsphase [37].

In der *Vorbereitungsphase* wählt ein Analyst eine Aufgabe aus und bestimmt dazu das Wissen, die Erfahrung und die Fähigkeiten der Nutzer, die die Aufgabe lösen sollen [37].

In der *Ausführungsphase* wird jedes Szenario beobachtet und die auftretenden Probleme werden protokolliert. [37].

#### 4.3.2 Nutzerbasierte Evaluationsmethoden

Bei der Entwicklung von interaktiven Software-Produkten stehen die Bedürfnisse der Nutzer im Vordergrund, daher ist es unerlässlich, dass auch echte Nutzer die Anwendung mit testen. Nutzerbasierte Evaluationsmethoden beinhalten eine Gruppe von Evaluationsmethoden, bei denen Nutzer ein Produkt evaluieren und deren Interaktionen protokolliert werden.

##### 4.3.2.1 Questionnaires

Questionnaires sind die derzeit am häufigsten verwendete Website-Usability Evaluationsmethode. Es ist einfacher und kostengünstiger als andere Ansätze um Benutzerfeedback zu sammeln [43]. Questionnaires hat eine systematische Herangehensweise zur Messung der Usability, durch vordefinierte Fragen und das eigene Berechnungssystem. Ein Fragebogen ist eine strukturierte Liste von Fragen, die verwendet werden, um Information von den Befragten zu sammeln [100]. Die Befragten haben die Möglichkeit ihr Feedback frei zu verfassen oder können vorgegebene Antworten verwenden [100].

Es gibt drei Typen von Fragebögen: Pre-Test-Fragebögen, Post-Task-Fragebögen und Post-Test-Fragebögen [43].

- *Pre-Test-Fragebogen* wird benutzt, um Informationen über die Teilnehmer zu sammeln, bevor sie mit dem Produkt interagieren. (welche Informationen?)
- *Post-Task-Fragebogen* wird benutzt, um unmittelbar nach jedem Szenario Feedback zu erhalten. (was für feedback?)
- *Post-Test-Fragebogen* wird benutzt, um ein gesamtes Feedback nach dem Test zu bekommen.

Innerhalb Post-Test-Fragebogen gibt es viele weitere Arten z. B. System Usability Scale (SUS), Questionnaire for User Interface Satisfaction (QUIS), Website Analysis and measurement inventory (WAMMI), Computer System Usability Questionnaire (CSUQ), Nielsen´s Attributes of Usability (NAU)...[100]. Aus der eigenen Recherche kamen SUS, QUIS und WAMMI Fragebögen häufiger zum Einsatz. Es hat vermutlich einen einfachen Grund, weil QUIS und WAMMI Fragebogen sich stärker auf Website-Usability beziehen und SUS einfach zu verwenden ist. Im Folgenden wird auf diese drei Arten näher eingegangen.

- *System Usability Scale (SUS)* wurde von John Brooke in 1986 entwickelt [100]. SUS ist ein kostengünstiges und effektives Instrument zum Messen von Usability [6]. SUS hat ein striktes Vorgehen und besteht aus zehn Fragen mit fünf Antwortmöglichkeiten zwischen 1 (stimmt überhaupt nicht) bis 5 (stimmt voll und ganz) [24]. Die Methode hat ihr eigenes Berechnungssystem. Von ungeraden Elementen wird 1 und von geradzahligen Elementen 5 subtrahiert. Die Ergebnisse der Fragen werden zusammen addiert und mit 2,5 multipliziert [24]. Der Score liegt zwischen 0 und 100. Je größer die Zahl desto besser schneidet die Usability einer Website ab.
- *Questionnaire for User Interface Satisfaction (QUIS)* wurde von Chin in 1988 von der Universität Maryland entwickelt [100], [6]. Das Ziel von QUIS ist die Zufriedenheit der Nutzer zu erfassen [6]. QUIS besteht aus 27 Aussagen, die in fünf Kategorien unterteilt sind: Reaktion auf Software, Bildschirm, Terminologie und Informationssystem, Lernen und die Systemfähigkeit [6]. Die Aussagen können auf einer Skala von 0 bis 9 bewertet werden, wobei 0 die niedrigste und 9 die höchste Bewertung ist [100].
- *Website analysis and measurement inventory (WAMMI)* wurde von Human Factors Research Group (HFRG) in 1999 entwickelt [100]. WAMMI besteht aus 20 standardisierten Fragen und basiert auf einer 5-Punkte-Skala zur Auswertung [100], [13]. WAMMI Fragen ermöglichen eine Analyse auf Basis von Webusability mit fünf Merkmalen: Attraktivität, Kontrolle, Effizienz, Nützlichkeit und Lernfähigkeit [100]. Die Anwendung von WAMMI ist wissenschaftlich von vielen Forschern bewiesen [13].

#### 4.3.2.2 Think-aloud

*Think-aloud* ist eine Methode, welche die Befragten ermutigt, ihre Meinung zu teilen [100]. Sie wurde von Ericsson und Simon entwickelt [68]. Die Nutzer sollen bei der Anwendung einzelner Funktionen laut denken. Die Think-aloud Methode hilft dabei zu erkennen, wie User die Informationen wahrnehmen während sie eine Aufgabe lösen [15]. Die gewonnenen Erkenntnisse sind nicht nur hilfreich zum Identifizieren von Usability-Problemen sondern auch als Anregung um neue Ideen zu entwickeln. Es gibt zwei Grundtypen von Think-aloud Methode: *Concurrent Think-Aloud* und *Retrospective Think-Aloud* [101] und eine neue Hybrid Methode [68].



- Bei *Concurrent Think-aloud (CTA)* (gleichzeitiges lautes Denken) werden die Teilnehmer gebeten, ihre Gedanken laut auszusprechen, während die Teilnehmer Aufgaben lösen [101].
- Bei *Retrospective Think-aloud (RTA)* werden die Gedanken nicht während des Lösen der Aufgaben laut ausgesprochen, sondern erst nachdem die Aufgaben gelöst sind [101].
- *Hybrid Methode (HB)* ist die Nutzung gleichzeitig beider Methoden CTA und RTA [68].

Einer Umfrage [101] nach ist die Concurrent Think-aloud die meist verwendet wird, mit 89% der 207 Befragten und lediglich nur 5% verwendeten die Retrospective Methode [101]. Warum die Variante Concurrent Think-Aloud so oft zum Einsatz kommt, gibt es dafür viele Gründe. CTA ist effizienter als die anderen zwei Varianten, weil wesentlich weniger Zeit pro Sitzung gebraucht wird, als bei RTA und HB [68]. Bei einem Test mit 20 Teilnehmern benötigt CTA nur 640 Minuten, RTA 1164 Minuten und HB 1233 Minuten [68]. Außerdem ist CTA für die Tester einfacher und verständlicher anzuwenden [101].

#### 4.3.2.3 Eye-Tracking

Eye-Tracking (ET) ist eine der fortgeschrittenen Methoden in Usability Testing [87] und ist eine bewährte Methode [54]. ET ist keine besonders neue Technologie. Bereits im Jahr 1879 fand Louis Emile Javal heraus, dass Texte ungleichmäßig von links nach rechts gelesen werden [78]. Während des Lesens stoppen oder bewegen die Leser ihre Augen über manche Stellen schneller als über andere. [45].

Heute wird ET in verschiedenen Bereichen eingesetzt z. B. in der Medizin bei Patienten, die in ihre motorischen Funktionen eingeschränkt sind, oder im Fußball um das Blickmuster eines Spielers wiederzugeben oder im Website Marketing. Eye-Tracking-Methode wird verwendet, um die Usability von Softwaresystemen zu bewerten und liefern Daten zum Verständnis der Fokuspunkte der Nutzer und liefert Daten zum Verständnis der Fokuspunkte der Benutzer in den Softwaresysteme [78]. Eye-Tracking beschreibt die Analyse von Augenbewegungen (z. B. Fixation, Sakkaden) [78] in einem statischen oder dynamischen Reiz (z. B. Bild oder Video). Fixation ist das gezielte Betrachten eines Objektes im Außenraum und Sakkaden sind schnelle Augenbewegungen [54], [78].

Um Blickverläufe bzw. Daten aufzeichnen zu können, werden Eye Tracker eingesetzt. Es gibt verschiedene Arten von Eye Tracker wie z. B. Eyetracking-Brillen oder monitorbasierter Eye Tracker. Diese funktionieren nach dem selben Prinzip und benötigen alle eine Lichtquelle [45]. Das Gerät sendet infrarot-nahes Licht ins Auge und wird von dem Auge reflektiert. Die Kamera erfasst das Auge des Nutzers und zeichnet das reflektierte Licht auf. Dadurch kann der Eye Tracker feststellen, wo der Nutzer hinschaut. Des

Weiteren wird Eye-Tracking Software benötigt [78] um die Daten auszuwerten.

Die Auswertung der Daten können in zwei Darstellungen visualisiert werden, als *Scanpath* und als *Heatmap*. Die Folgende Grafik (Abbildung 4.2) zeigt ein Beispiel der der beiden Darstellungen. Linke Seite ist der Scanpath und rechte Seite ist die Heatmap [57].

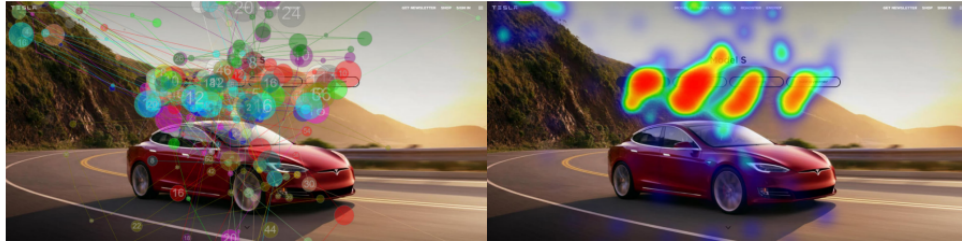


Abbildung 4.2: Visualisierung der Daten als Scanpath und Heatmap

Wenn der User eine Website besucht, versucht er die Website zu scannen und wird nicht Wort für Wort lesen. Eye Tracking ermöglicht es das Surfverhalten der Nutzer aufzuzeigen, wie und welche Elemente einer Website die Nutzer betrachten. Dabei sind laut der Studie [70] vier *Haupt-Scannenmuster* festzustellen.

- *Das F-Pattern*: Das F-förmige Scanmuster ist gekennzeichnet durch viele Fixierungen, die sich oben und links auf der Seite konzentrieren. Das Leseverhalten vieler Menschen ist von links nach rechts und die meisten Menschen lesen beim Besuch einer Website nicht Wort für Wort, sondern überfliegen die Website um sich einen Überblick zu verschaffen. In einer Studie aus dem Jahr 2008 fand die NNGroup heraus, dass "auf einer durchschnittlichen Website die Besucher maximal 28% der Wörter während eines durchschnittlichen Besuchs lesen." Daraus lässt sich ableiten, dass Besucher die Website zu Beginn flüchtig betrachten. Weist eine Website dieses Lesemuster auf, ist es auch ein Indiz dafür, dass das Design bzw. die Inhalte einer Website noch nicht ausgereift sind, weil offenbar Abschnitte beim Lesen übersprungen werden [70].
- *Das spotted Pattern* beinhaltet die Fixierung auf bestimmte Wörter oder Wortgruppen, die über die ganze Seite verteilt sind. Es gibt zwei Gründe [70], warum dieses Muster entsteht. Zum einen werden diese Wörter durch Bulletpoints oder mit einer Farbe versehen. Zum anderen sind es genau die Wörter, wonach die Nutzer suchen [70].
- *Das Layer-Cake Scanning-Pattern*: Dieses Muster ähnelt einem Kuchen mit mehreren Schichten. Die Nutzer achten auf Überschriften und lesen manche Abschnitte vollständig durch andere aber nicht [70].
- *Das Commitment Pattern*: Es handelt sich nicht mehr um das Scannen einer Webseite, bei dem der Nutzer die Seite überfliegt, sondern um das



tatsächliche Lesen [70]. Der Nutzer nimmt bei diesem Muster im Vergleich zu anderen Mustern mehr von der Webseite wahr und verbringt auch hier mehr Zeit. Da eine lange Verweildauer zu den Attributen gehört, welche eine gute Webseite ausmachen, ist es ein Indiz dafür, dass eine solche Webseite für den Nutzer interessant erscheint.

#### 4.3.2.4 Interview

Bei Interview werden Nutzer zu einem bestimmten Ereignis befragt. Interviews sind eine gute Möglichkeit um tiefgreifende Daten von einer kleinen Anzahl der Befragten zu erhalten [100]. Eine Studie möchte mithilfe von Interviews mehr Informationen über Usability oder mehr von der Erfahrung der User beim Benutzen der Anwendung erfahren [69]. Eine weitere Studie benutzt Interviews um die Daten, die durch Eye-Tracking erhoben wurden, zu verifizieren [78]. Eine andere Studie verwendet Interviews um weitere Details zum vorher durchgeführten Usability-Test zu erhalten [65]. Es fällt auf, dass bei keiner der Studien [69], [78], [65] ausschließlich Interviews zur Usability-Evaluation verwendet wurden, sondern immer nur als Ergänzung zu anderen Methoden um mehr Details über die Nutzer nach Usability-Test zu erfahren.

#### 4.3.2.5 Usability-Testing

*Usability-Testing* wird mit ein paar anderen Begriffen synonym verwendet wie z. B. Usability-Test, User-Test, User-Testing [42]. Usability-Testing kann für verschiedene Arten von Designs genutzt werden, von der Benutzeroberfläche bis zum physischen Produkt [6]. Bei Usability-Testing werden direkte Beobachtungen und Gespräche mit potenziellen Benutzern eines Produktes durchgeführt, um zu sehen, wie sie mit dem Produkt interagieren [20]. Solche Tests sind der einfachste und der grundlegendste Usability-Evaluierungsansatz [3].

#### 4.3.2.6 Remote Usability Testing

*Remote Usability Testing* ist eine experimentelle Methode um die Usability einer Website zu messen [80]. Die Methode ermöglicht Forschern die Usability von Webseiten durch das sammeln der Informationen von Nutzern, die von verschiedenen Orten aus teilnehmen können, zu evaluieren [85], [94]. Um den Test durchzuführen, werden Aufgabenszenarien erstellt, die die Funktionalität der Software wiedergeben. Tester werden gebeten, die Aufgaben zu lösen und dabei wird die Zeit gemessen, wie lange die Tester zum Lösen der jeweiligen Aufgabe benötigen [80].

#### 4.3.3 Toolbasierte Evaluationsmethode

Toolbasierte UEM sind neue Wege um Websites zu verbessern [32]. Toolbasierte UEM sind Methoden, bei denen Tools zum Einsatz kommen, anstatt

Nutzer oder Experten einzubeziehen. Tools haben den Vorteil, dass sie die Usability objektiv evaluieren und frei von subjektiven Einflüssen sind.

#### 4.3.3.1 Klick-Analyse

Klick Analyse ist eine auf Webanalyse basierende Technik, die verfolgt, wo die Benutzer auf die Website klicken [32]. Klick Analyse wird genutzt, um die Zufriedenheit der User und die Konversionsrate von Webseiten zu verbessern [32]. Zufriedenheit ist eine der Metriken, die eine gelungene Usability auszeichnet, daher eignet sich die Klick-Analyse auch zur Usability-Evaluation. Die Klick-Analyse ist im Bereich der Web-Usability eine Methode, mit der die Nutzer und ihr Nutzungsverhalten analysiert werden können. Das Nutzungsverhalten wird in Form von Heatmaps (Wärmebild) visuell dargestellt. Das Wärmebild ist im Zusammenhang mit der Wärmebildkamera bekannt. Die Abstufung der Farben von rot bis blau stellt die Abstufung der Temperaturen dar. Analog dazu stehen die unterschiedlichen Farben der Heatmap für die unterschiedlichen Intensivitäten der Nutzerinteraktion. Die unterschiedlichen Farben kommen dadurch zustande, wie oft auf ein Element einer Seite geklickt wird. Das bedeutet, je häufiger Aktivitäten in einem Bereich stattfinden, desto wärmer ist die Farbe.

Es gibt derzeit drei Haupttypen von Heatmaps (Abbildung 4.3), die in verschiedene Tools auf dem Markt angeboten werden: Klick-Heatmaps, Move-Heatmaps und die Scroll-Heatmaps. [53] [107].

Die häufig benutzte Heatmap ist die durch Klicks erzeugte Heatmap (Abbildung 4.3 a) [53]. Diese zeigt an, wo und wie oft die Nutzer auf einen Bereich einer Seite klicken (Abbildung 4.4). Wie auf der Abbildung zu erkennen ist, werden nicht nur Klicks auf klickbare Elemente aufgezeichnet, sondern auch Bereiche, bei denen Klicks keine Aktionen auslösen.

Der zweite Heatmaps Typ ist die Move-Heatmap (Abbildung 4.3 b). Sie gibt den Mausbewegungsverlauf wieder, wie sich der User mit der Maus über einem bestimmten Bereich bewegt.

Der dritte Typ ist die Scroll-Heatmap (Abbildung 4.3 c). Die Scroll-Heatmap ist ein Maß für die Scrolltiefe. Sie gibt Auskunft darüber, wie viele Nutzer bis zu welcher Stelle einer Seite scrollen.

Es ist denkbar einfach, wie die Daten für die Heatmap-Analyse erhoben werden. Nach erfolgreicher Anmeldung für ein Tool wird ein Tracking-Script erzeugt und dies kann in den <head> oder <body> Tag einer Website eingebunden werden, abhängig von der Empfehlung des ansprechenden Anbieters [32].

Die Daten werden nicht nur in Form von Heatmaps dargestellt, sondern wahlweise auch als Overlay, Confetti View und List [32].

- Der *Overlay* Bericht zeigt die genaue Anzahl an Klicks überlagert auf einem Element einer Seite. Der Bericht stellt nicht nur die Anzahl der Klicks zur Verfügung, sondern auch Information über das Land, die Browser, den Gerätetyp oder das Betriebssystem der Nutzer.

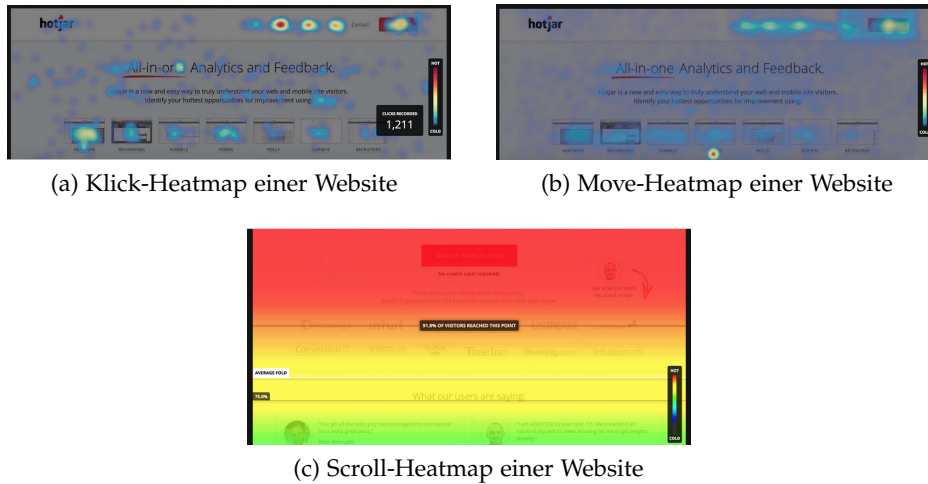


Abbildung 4.3: Heatmaps

- In dem *Confetti View* Bericht werden die Klicks als Punkte angezeigt. Jeder Punkt hat eine bestimmte Farbe als Darstellung von Information wie Browser, Gerätetyp oder Land.
- In dem *List* Bericht werden die Daten in Form von einer Liste dargestellt. Die Liste ähnelt der Darstellung einer Tabelle und hat mehrere Spalten, in denen die Elementen mit der zugehörigen Klickanzahl aufgelistet werden.

Die unterschiedlichen Darstellungen der Klick-Analyse liefern somit wertvolle Information über die Nutzer sowie deren Nutzerverhalten. Damit können Maßnahmen abgeleitet werden um die Zufriedenheit der Nutzer zu verbessern.

#### 4.3.3.2 Automatisierte Tools

Es kann einige Faktoren geben, die die Benutzerfreundlichkeit der Website beeinflussen. Dazu gehören beispielsweise lange Ladezeit, defekte Links oder nicht vorhandene allgemeine Suchoptionen. [26]. Automatisierte Tools analysieren, ob die Benutzeroberfläche den Usability-Standards entspricht [85]. Es prüft die Qualität des HTML-Codes gemäß den Richtlinien [85]. Mit Hilfe eines Usability-Modells lässt sich die Qualität der Usability einer Website in Zahlen ausdrücken [9].

Bei expertenbasierte und nutzerbasierte UEM ist es bei manchen Methoden nicht möglich oder schwer die Usability einer Webseite zu messen und die Ergebnisse in Zahlen darzustellen z. B. Heutistische Evaluation der Interviews.... Dieses Problem tritt bei automatischen Analyse Tools nicht auf. Automatische Analyse Tools sind Tools zur Messung der Qualität von Webseiten. Jakob Nielsen misst Qualität einer Webapplikation zu bewerten mit Hilfe von fünf Metriken: Erlernbarkeit (learnability), Effizienz (efficiency), Einprägsamkeit (memorability), Fehler (errors) und Zufriedenheit (satisfaction) [63]. Um diese Metriken messbar zu machen, müssen Eigenschaften

einer Website diesen Metriken zugeordnet werden z. B. die Ladegeschwindigkeit unter Effizienz oder die URL unter Einprägsamkeit. Die Studie [9] schlägt ein mögliches Usability-Modell (Abbildung 4.4) für die automatische Analyse vor.

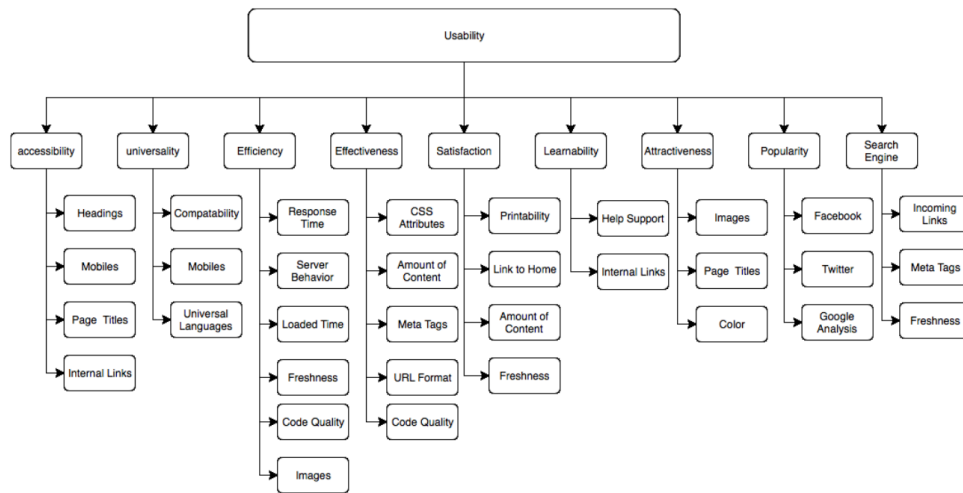


Abbildung 4.4: beispielhaftes Usability Modell

Die automatische Toolanalyse ist die Analyse von HTML-Code. Sie verifiziert, ob die Qualität des HTML-Codes gemäß der Richtlinien [85] z. B. die Qualitätsstandards des W3C [9], [99], World Wide Web Consortium. W3C ist eine internationale Gemeinschaft, die Standards entwickelt um das langfristige Wachstum des Webs zu sichern.

Bei einem Blick auf das Usability-Modell fällt auf, dass beispielsweise "Meta Tags" sowohl unter "Effectiveness" als auch unter "Search Engine" und "Images" unter "Efficiency" und "Attractiveness" zugeordnet werden. Das muss bei der Auswertung des Ergebnisses berücksichtigt werden.

#### 4.4 ZUSAMMENFASSUNG

Damit ein Grundverständnis für das Thema rund um den Begriff *Usability* geschaffen wird, wurde im ersten Schritt mehrere Begriffe wie Usability, Usability-Metriken, Usability-Probleme, Usability-Evaluationsmethoden näher erläutert. Aufgrund des breiten Spektrums an Usability-Evaluationsmethoden wurden die Methoden nach Kategorien klassifiziert. Es gibt mehrere Merkmale, wonach die Methoden klassifiziert werden können wie z. B. nach dem Untersuchungszweck oder nach der Analyseform. Für die Arbeit werden die Methoden nach Gutachter/Evaluator klassifiziert. Dabei unterscheiden sich drei Kategorien voneinander: expertenbasierte UEMs, nutzerbasierte UEMs und toolbasierte UEMs. Im letzten Schritt werden die Methoden aus der Recherche der entsprechenden Kategorie zugeordnet und vorgestellt. Zu den expertenbasierten UEMs gehören die Heuristische Evaluation und das Cognitive Walkthrough. Die nutzerbasierten UEMs beinhalten die folgenden Methoden: Questionnaire, Think-aloud, Eye-Tracking, Interview,

Usability-Testing und Remote Usability-Testing. Zu der dritten Kategorie, den toolbasierten UEMS zählen die Klick-Analyse und die Automatisierten Tools.

## VERGLEICHE DER METHODEN

---

In diesem Kapitel werden die Methoden miteinander verglichen. Der Vergleich hilft dabei, die Methoden besser zu verstehen und um den passenden Anwendungsfall für die jeweilige Methode zu erkennen. Außerdem dient der Vergleich als Grundlage für die Erstellung eines Schemas zur Auswahl der Usability-Evaluationsmethoden. Es werden UEMs aus der eigenen Recherche verglichen, die von 2015 bis 2021 am meisten verwendet wurden. Die UEMs sind: Questionnaires, Heuristische Evaluation, Automatisierte Tools, Think-aloud, Eye-Tracking, Interview, Usability Testing, Cognitive Walkthrough und Klick-Analyse. Remote Usability Testing wird nicht in dem Vergleich berücksichtigt, weil die Vorgehensweise dem normalen Usability-Testing ähnelt, bis auf den großen Unterschied, dass der Test remote stattfindet.

Bevor ein Vergleich stattfinden kann, müssen die Vergleichskriterien bestimmt werden. Die Vergleichskriterien sind dazu da, um die Methoden systematisch zu charakterisieren und umfänglich zu vergleichen.

### 5.1 BESTIMMUNG DER VERGLEICHSKRITERIEN FÜR UEM

In dieser Arbeit dreht es sich um Methoden zur Usability-Evaluation von Websites. Aus dem Grund beziehen sich die Vergleichskriterien auf folgende Punkte, die unmittelbar mit einer Evaluation im Zusammenhang stehen: (1) Methoden, (2) Usability, (3) Websites, (4) Daten, (5) Ressourcen und (6) Stärke und Schwächen.

(1) Mit dem Punkt *Methode* ist damit gemeint, was für ein *Typ* die Evaluationsmethode ist. Handelt es sich um eine Inspektionsmethode, Testmethode oder eine Befragung.

- *Inspektionsmethode*: Im Kontext Usability sind Inspektionsmethoden formale Methode zur Überprüfung der Benutzeroberfläche durch Evaluatoren. Usability Inspektion strebt an, Usability Probleme am Design aufzudecken [62]. Für die Inspektion einer Benutzeroberfläche ist es nicht notwendig, dass sie fertig implementiert sein müssen. Daher können solche Inspektionsmethoden in der frühen Stufe des Entwicklungsprozess angewendet werden [62]. Daher erfordert solche Methoden Spezialisten wie z. B. Designer, Software Entwickler oder Usability-Experten um die Benutzeroberfläche zu beurteilen nach etablierten Usability-Prinzipien [100].
- *Testmethode*: Solche Methoden helfen zu verstehen, wie reale Nutzer bestimmte Software bedienen. Während Usability-Tests Nutzer ein Sys-

tem oder einen Prototyp verwenden, um eine vorgegebene Reihe von Aufgaben zu erledigen, oder die Ergebnisse der Teilnahme werden aufgezeichnet [29].

- *Befragung*: Befragungen holen Feedback von Nutzern und werden häufig bei Usability-Tests eingesetzt. Der Fokus liegt nicht auf Untersuchung bestimmter Aufgaben oder Leistungsmessung, der Ziel ist vielmehr subjektive Eindrücke zu sammeln [29].

(2) In dem vorherigen Kapitel wurde der Begriff Usability näher erklärt, was unter dem Begriff Usability zu verstehen ist. Die Usability wird anhand von den Usability-Metriken gemessen. Diese müssen berücksichtigt werden, wenn es darum geht, die Usability zu verbessern. Die Metriken sind: *Effektivität, Effizienz, Zufriedenheit, Erlernbarkeit, Einprägsamkeit und Fehler*. Das Vergleichskriterium ist somit Usability-Metriken.

- *Usability-Metriken*: Effektivität, Effizienz, Zufriedenheit, Erlernbarkeit, Einprägsamkeit und Fehler

(3) Websites sind Software-Produkte. Jede Software durchläuft während der Entwicklung mehrere Phasen. In der Softwareentwicklung wird von einem *Software Development Lifecycle* gesprochen. Der Software Lebenszyklus *Software development lifecycle (SDLC)* beschreibt den Lebenszyklus eines Software-Produkts von der Schöpfung bis zum Einstellen der Software. Die SDLC lässt sich in mehreren Phasen gliedern: Anforderungsanalyse, Design, Entwicklung, Testen, Produktion und Wartung[82]. Zur Vereinfachung wird für den Vergleich die SDLC in vier Phasen unterteilt:

- *Vor der Implementierung*: Diese Phase beinhaltet die Anforderungsanalyse und die Designphase
- *Implementierung*: In dieser Phase werden die Anforderungen umgesetzt.
- *Testen*: Nach der Entwicklungsphase wird die Software auf ihre Funktionalität anhand der vordefinierten Anforderungen getestet.
- *Betrieb*: Das ist die Phase, in der das Produkt bereits in Produktion ist und gewartet wird.

(4) Zu jeder Evaluation werden die *Daten* erhoben um sie später auswerten zu können. Da stellt sich die Frage, welche Arten von Daten gesammelt werden. Es werden zwischen quantitativ und qualitative Daten unterschieden.

- *quantitative Daten*: Qualitative Methoden werden in Situationen eingesetzt, in denen das Ergebnis nicht offensichtlich ist oder das sich das Ergebnis nicht für statistische Analysen eignet [61].
- *qualitative Daten*: Quantitative Methoden untersuchen die Auswirkungen bestimmter Umstände auf ein Resultat, welches in Zahlen ausgedrückt werden kann[61], z. B Anzahl der Besucher, Anzahl der neuen Besucher oder Verweildauer.



(5) *Ressourcen* spielen eine wichtige Rolle bei der Evaluation. Da die Ressourcen nicht unbegrenzt sind, muss geschaut werden, wie viele Ressourcen zur Verfügung stehen. Zu den Ressourcen gehören die Zeit, materielle Ressourcen und personelle Ressourcen.

- *Ressourcen*: Zeit, materielle Ressourcen und personelle Ressourcen.

(6) Jede Methode hat ihre *Stärken* und *Schwächen* und diese sollen bekannt sein, damit von Situation zu Situation die richtige Methode gewählt werden kann.

- *Stärken*
- *Schwächen*

Alle Punkte im Zusammenhang mit der Evaluation sind einbezogen worden. Die eben ermittelten Vergleichskriterien dienen als Merkmale für den Vergleich der Usability-Evaluationsmethoden.

## 5.2 VERGLEICH DER USABILITY-EVALUATIONSMETHODEN

In diesem großen Abschnitt werden die Evaluationsmethoden in der gleichen Kategorie miteinander verglichen. Der Vergleich orientiert sich an den Vergleichskriterien (1)-(6) aus dem vorigen Abschnitt. Nach jedem Vergleich einer Kategorie werden die Ergebnisse in einer Tabelle zusammengefasst.

### 5.2.1 Vergleich der expertenbasierten UEM

- (1) *Methodentyp*: Heuristische Evaluation (HE) ist eine Inspektionmethode, bei der die Usability anhand von Usability Prinzipien untersucht wird [40], beispielsweise die Prinzipien von Nielsen.

CW ist eine Inspektionmethode [11], da bei dieser Methode Experten die Usability inspizieren. Jedoch kann CW auch eine Testmethode [37] sein, weil die Usability auch getestet wird, indem eine Reihe von Aufgaben gelöst werden. Dies hat den Vorteil, dass die Bedienung der Anwendung simuliert wird [88].

- (2) *Usability-Metriken*: Einer Studie nach [2] kann HE Effektivität, Effizienz und Zufriedenheit eines Benutzers steigern. Die Ergebnisse einer weiteren Studie zeigten auch, dass HE stärker auf das Identifizieren von Usability-Problemen in Bezug auf Effektivität und Zufriedenheit fokussiert ist [51]. Jedoch im Vergleich zu CW ist HE bei allen anderen Usability-Problemen in Bezug auf die sechs Usability-Metriken überlegen [51].

Bezüglich der Usability-Metriken findet CW mehr Effektivität- und Erlernbarkeit-Probleme [51].



- (3) *Software development lifecycle*: Cognitive Walkthrough kann während aller Phasen der Software Entwicklung eingesetzt werden [11]. Außerdem versuchen Evaluatoren während des CW Aktionen auszuführen, die für gewöhnliche Benutzer schwierig sein können, indem sie versuchen die Schnittstellen zu verstehen und ihre potenzielle Auswirkungen auf die Benutzer [51]. Daher kann CW in der frühen Entwicklungsphase durchgeführt werden, um die Anforderung der Benutzer zu erfüllen [51].

Genau so wie CW kann auch Heuristische Evaluation jederzeit während der Entwicklung einer Software durchgeführt werden [40], doch es wird empfohlen im frühen Entwicklungszyklus zur Identifizierung kritischer Usability-Probleme anzuwenden [18].

- (4) *Art der Daten*: Aufgrund der Tatsache, dass sowohl bei HE als auch CW die Daten empirisch erhoben werden und diese beispielsweise textuell, verball oder in ähnlicher Form wieder gegeben werden, sind beide Methoden qualitative Daten.
- (5) *Ressourcen*: Wenn die Zeit begrenzt zum Evaluieren ist, kann HE eine gute Lösung sein. Nielsen fand heraus, dass bei HE drei bis fünf Evaluatoren schon einen Großteil der schwerwiegendsten Usability-Probleme finden können [18], [95]. Damit ist HE eine kostengünstige Methode und kann eine hohe Anzahl an Usability Probleme identifizieren [51].

In der Studie [51] nahmen fünf Evaluatoren teil, um ein System zu inspizieren nach HE und CW. Das Ergebnis zeigt: HE identifizierte deutlich mehr Usability Probleme als CW ( 317 zu 104) bei gleichen Anzahl an Evaluatoren. Deswegen ist HE kosteneffektiver als CW.

- (6) *Stärken*: Heuristische Evaluation hat den Vorteil, dass mit nur drei bis fünf Experten bereits ca. 75% der Usability-Probleme erkannt werden. Es bedarf kein besonderes Equipment um HE durchzuführen. Dadurch, dass diese Methode mit einer geringen Anzahl von Experten auskommt, erfolgt die Auswertung der Ergebnisse dementsprechend auch zügig und nimmt nicht viel Zeit in Anspruch. Die HE gilt als eine schnelle und günstige Methode [88] um Feedback einzuholen.
- (6) *Schwächen*: Die HE generiert Feedback nur aus der Sicht der Experten und bezieht keinerlei echte Nutzer zu der Analyse mit ein. So ist es nicht möglich, alle Usability-Probleme zu aufzudecken. Eine Studie [5] fand heraus, Probleme, welche durch Heuristische Evaluation identifiziert sind, sind nicht die gleichen Probleme, die durch Usability Testing gefunden werden und das in 40% (34 von 86) der Fällen. Der Erfolg einer Heuristischen Evaluation ist von den ausführenden Experten abhängig, weil erfahrene Evaluatoren mehr Probleme erkennen können, als unerfahrene Evaluatoren [51], [88].

—	HEURISTISCHE EVALUATION	COGNITIVE WALKTHROUGH
METHODENTYP	Inspektion	Inspektion, Test
USABILITY-METRIKEN	Effektivität, Effizienz, Zufriedenheit, Erlernbarkeit, Einprägsamkeit, Fehler	Erlernbarkeit, Einprägsamkeit
SDLC	alle Phasen	alle Phasen
ART DER DATEN	qualitative	qualitative
RESSOURCEN	schnell, kostengünstig, 3-5 Experten	3-5 Experten
STÄRKEN	geringe Evaluatoren, Identifizierung hoher Anzahl an Usability-Problemen	fokussiert auf die einfache Erlernbarkeit eines Systems - simuliert die Bedienung echter Nutzer
SCHWÄCHEN	Ergebnisse anhängig wie erfahren die Experten sind	findet wesentlich weniger Probleme als HE

Tabelle 5.1: Zusammenfassung des Vergleichs expertenbasierter Methoden

### 5.2.2 Vergleich der nutzerbasierten UEM

- (1) *Methodentyp*: Mit Questionnaires und Interviews werden Nutzer zu einem bestimmten Thema befragt.

ET ist eine Testmethode, weil die Teilnehmer die Anwendung testen, indem sie versuchen vorgegebene Aufgaben zu lösen.

Bei Usability-Testing bekommen User ausreichend Zeit um sich mit der Software vertraut zu machen, bevor sie mit dem Test der Software beginnen [97]. Damit ist diese Methode auch wie ET eine Testing Methode.

Think-aloud ist sowohl eine Befragungsmethode als auch eine Testmethode. Bei Think-aloud werden die Teilnehmer gebeten, ihre Gedanken laut ausszusprechen während sie Aufgaben lösen [46].

- (2) *Usability-Metriken*: Die WAMMI Questionnaires adressieren u. a. folgende Usability Metriken: Attraktivität, Effizienz und Einprägsamkeit [100] und SUS Questionnaires u.a. die Usability Metriken: Effektivität, Zufriedenheit, Erlernbarkeit und Fehler [48].

TA ist in der Lage Daten zu Nutzer Zufriedenheit zu sammeln, dabei spielt es keine große Rolle, um welche Art von TA es sich handelt [67], [68].

In ET wird die Anzahl der Fixierungen benutzt um die Effizienz (wie schnell das Ziel erreicht werden?) und Effektivität (wird das Ziel vollständig bzw. korrekt erreicht?) von Software zu untersuchen [78], [108].

In keine der bisher untersuchten Studien wird das Interview als einzige Methode benutzt, um die Usability zu evaluieren. Interview wurde vor allem zusammen mit anderen Methoden kombiniert [60], [69], [78], [1]. Es wurde auch nicht explizit thematisiert, welche Usability-Metriken Interviews verbessern sollen. Eine Studie [69] verwendet Usability-Testing und Interview mit dem Ziel die Effektivität, Lernfähigkeit, Einprägsamkeit, Effizient, Fehler und Zufriedenheit der Usability einer Website zu verbessern.

Usability-Testing weist den grundlegendsten Usability-Evaluierungsansatz auf [3]. Daher ist diese Methode flexibel. Je nach Anwendungsfall sind die Aufgaben, Fragen oder das Setup eines Tests unterschiedlich. Diese Flexibilität hat den Vorteil, dass mehrere Metriken berücksichtigt werden. So ist Usability Testing beispielsweise in der Lage auf Effektivität, Effizienz, Erlernbarkeit und Zufriedenheit zu testen [33].

- (3) *SDLC*: Basierend auf den Rechercheergebnissen ist festzustellen, dass dieser Punkt in den Studien nicht thematisiert oder nicht gezielt darauf eingegangen wurde. Es fällt jedoch auf, dass die Kategorie nutzerbasierte UEMs fast ausschließlich auf Anwendungen, die bereits im Betrieb sind, angewendet werden. Daher eignen sie sich auf jeden Fall für die Betriebsphase.

Questionnaires richten sich an echte Nutzer, daher ist es sinnvoll, sie in der Betriebsphase einzusetzen.

Die Studien [76], [54], [81], [30] verwenden ET für Anwendungen, die bereits im Betrieb sind oder in der Testphase [79].

Usability Testing wird angewendet, wenn potenzielle oder echte Nutzer die Software testen. Daher eignet sich die Methode für die Testphase und für die Betriebsphase.

In den Studien [65], [60], [69], [1] werden Interviews nach den Usability-Tests durchgeführt, daraus lässt sich schließen, dass Interviews in der Testphase und Betriebsphase zum Einsatz kommt.

- (4) *Art der Daten*: Einer Studie nach können mit Fragebögen quantitative Daten gesammelt werden und qualitative Daten mit Usability Testing [47]. Meiner Meinung nach können mit Fragebögen auch qualitative Daten erhoben werden, je nachdem wie die Fragen formuliert sind.

TA ist in der Lage sowohl qualitative als auch quantitative Usability-Probleme [14], [68] zu erkennen.

Die Mehrheit der ET-Studien konzentrieren sich auf das Sammeln quantitativer Informationen über den Blickpunkt und die Verweildauer beim Betrachten eines Objekts [78], [81].

Interviews dienen zur qualitativen Analyse, weil die Nutzer nach ihren Erfahrungen und Meinungen befragt werden, nachdem sie die Anwendung getestet haben.

Beim Usability-Testing werden die Teilnehmer von Gutachtern beobachtet [83]. Diese notieren die Beobachtungen und generieren daraus qualitative Ergebnissen. Des Weiteren kann Usability-Testing auch quantitative Ergebnisse liefern, wenn beispielsweise die Zeit gemessen wird, wie lange die Nutzer braucht um bestimmte Aufgaben zu lösen [4].

- (5) *Ressourcen*: Questionnaires sind leichter und kostengünstiger als andere Ansätze um Feedback von Nutzern einzuholen [100]. Jedoch sollte berücksichtigt werden, dass die User die an der Befragung teilnehmen, nicht unbedingt die Gesamtheit aller Nutzer spiegelt [98].

Um statistische Aussagekraft zu erhalten wird bei Think-aloud eine Stichprobengröße von 20 Teilnehmern empfohlen [67]. In dieser Studie hat TA im Durchschnitt ca. 28 Stunden gebraucht von der Ausführung bis zur Analyse der Daten. TA ist daher auf den ersten Blick ein langwieriger Prozess, jedoch auf 20 Teilnehmer runter gerechnet, wird nur ca. 1,5 Stunden pro Teilnehmer gebraucht. Das ist meiner Meinung nach ein vertretbarer Zeitaufwand.

Nielsen empfiehlt 46 Teilnehmer [52] für Eye-Tracking, um aussagekräftige Ergebnisse zu bekommen. Der Einsatz von ET ist damit sehr zeitintensiv. Außerdem müssen sich Anwender von ET darüber bewusst sein, dass der Einsatz von ET mit sehr hohen Kosten verbunden ist.

Das Setup von ET beinhaltet Hardware (Eye-Tracker, PC sowie zwei Monitore), Software, Personal zum Durchführen von ET-Experimente, Zeit zum Ausführen so wie Zeit für die Analyse der Daten [49].

Die Studien zeigen [65], [60], [69], [1], dass Interviews durchgeführt werden, um mehr Details zu erfahren, die anders nicht schwer zu bekommen sind. Außerdem wurden bei den Studien immer nur eine geringe Anzahl von Teilnehmern, ca. 5-6 Personen interviewed. Die Vermutung liegt nahe, dass Interviews zeitaufwändig ist.

Qualitative Methoden bei nutzerbasierten Methoden haben manchmal die Schwäche, dass bei kleinen Teilnehmerzahl die Ergebnisse zu verallgemeinern [39]. Der Studie [37] nach benötigt Usability-Testing keine besonderen Anforderung oder hohe Kosten. Weiter gibt die Studie bekannt, dass ein Tester bis zu 25% der Usability-Probleme identifizieren kann und bei 15 Testern (Abbildung 5.1) sogar 100% der Usability-Probleme [37].

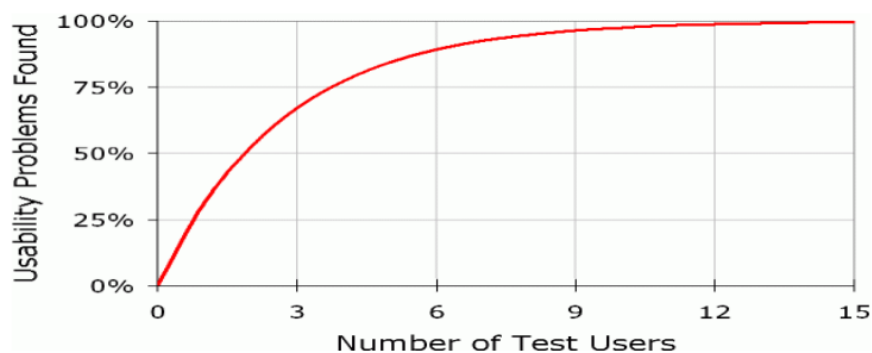


Abbildung 5.1: Empfohlene Anzahl der Tester (5-15) (Nielsen 2000) [78]

- (6) *Stärken*: Questionnaires sind leicht und kostengünstig umzusetzen [100]. Des Weiteren haben Questionnaires ihr eigenes Berechnungssystem um die Usability nach striktem Vorgehen zu bewerten [100]. Ein weiterer Vorteil von Questionnaires ist, dass wenn die Fragebögen einmal erstellt sind, können eine große Anzahl Nutzer daran teilnehmen. So haben die Studien [90], [86] Fragebögen an 516 und 550 potenzielle Teilnehmer versendet.

TA hat die große Stärke, die besten Erkenntnisse in den kognitiven Prozessen der Endnutzer zu gewinnen [55], da Nutzer beim Testen der Anwendung laut ihre Gedanken aussprechen.

ET ist in der Lage, das Verhaltensmuster von Nutzern wiederzugeben und das Areas of Interest (AOI) kann identifiziert werden [78]. Darüber hinaus kann ET auch Informationen über die Ursachen von Usability-Problemen liefern. [4].

Die Durchführung eines Interviews ist flexibel. Die Interviewfragen können für jeden Einzelnen Befragten maßgeschneidert erstellt werden [65], um unterschiedliche Erfahrungen jedes Einzelnen zu sammeln.

Usability Testing ist eine effektive Methode zur Evaluation von Benutzeroberfläche [4]. Diese Methode liefert Informationen von echten Nutzern und hilft Erwartungen von Kunden zu verstehen, wie das Produkt verbessert werden kann [4].

- (6) *Schwächen*: Die Studie [86] hat zwar Fragebögen an 550 potenzielle Teilnehmer versendet, jedoch wurden nur rund 33% Fragebögen zurück gemeldet. Dieser Aspekt muss bei der Auswertung mitberücksichtigt werden.

Eine Untersuchung in Bezug auf die Beziehung zwischen Augenbewegungen und TA-Protokolle zeigen, dass die verbale Bericht von TA unvollständig sein können [66]. Es kann daran liegen, dass Tester zwei Aufgaben zur selben Zeit erledigen müssen, sowohl das Testen einer Anwendung als auch das lautstarke Denken. TA Testsitzungen werden in einem laborbasierten Umfeld durchgeführt [67]. Das hat zur Folge, dass dieses Setup nicht die natürliche Umgebung der Nutzer wieder und daher ist es möglich, dass das Verhalten der Nutzer nicht vollständig erfasst wird [67].

—	QUESTION-NAIRES	THINK-ALOUD	EYE-TRACKING
METHODENTYP	Befragung	Testmethode, Befragung	Testmethode
USABILITY-METRIKEN	Effizienz, Einprägsamkeit, Effektivität, Zufriedenheit, Erlernbarkeit und Fehler	Zufriedenheit	Effizienz und Effektivität
SDLC	Betriebsphase	Betriebsphase	Testphase, Betriebsphase
ART DER DATEN	quantitativ, qualitativ	quantitativ, qualitativ	qualitativ
RESSOURCEN	kostengünstig	vertretbarer Zeitaufwand, 20 Teilnehmer	zeitintensive, hohe Kosten, mindestens 46 Teilnehmer
STÄRKEN	einfach umzusetzen, große Teilnehmerzahl möglich	wertvolle Erkenntnisse in kognitiven Prozessen der Nutzer	identifiziert von Verhaltensmustern und AOI, liefert Information über Ursachen von Usability-Problemen
SCHWÄCHEN	potenzielle geringe Rückmeldung	laborbasierte Umgebung, unvollständige Information	—

Tabelle 5.2: Zusammenfassung des Vergleichs nutzerbasierter Methoden (1)

	INTERVIEWS	USABILITY TESTING
METHODEN-TYP	Befragung	Inspektion, Test
USABILITY-METRIKEN	Effektivität, Lernfähigkeit, Einprägsamkeit, Effizienz, Fehler und Zufriedenheit	Effektivität, Effizienz, Erlernbarkeit und Zufriedenheit
SDLC	Testphase, Betriebsphase	Testphase, Betriebsphase
ART DER DATEN	qualitativ	qualitativ, quantitativ
RESSOURCEN	zeitaufwändig, geringe Teilnehmer (5-6)	keine hohen Kosten, 5-15 Tester
STÄRKEN	flexible Durchführung, tiefgreifende Erfahrungen	hilft Erwartungen von Nutzern zu verstehen, effektive Methode
SCHWÄCHEN	zeitaufwändig	–

Tabelle 5.3: Zusammenfassung des Vergleichs nutzerbasierter Methoden (2)

### 5.2.3 Vergleich der toolbasierten UEM

- (1) *Methodentyp*: Klick-Analyse ist eine Testmethode [32]. Die Ergebnisse oder Beobachtungen aus der Klick-Analyse führen dazu, dass Änderungen auf der Webseite vorgenommen werden können und diese verschiedenen Änderungen müssen getestet werden [32] um feststellen zu können, ob eine Verbesserung der Usability auftritt.

Automatisierte Tools ist auch eine Testmethode. Die Tools führen automatisch Tests auf Websites durch.

- (2) *Usability-Metriken*: Klick-Analyse kann die Bereiche wiedergeben, die am meisten Aufmerksamkeit der Nutzer bekommen. Die Methode kann beispielsweise identifizieren, wie hoch die Klickanzahl der Elemente auf der Website ist oder redundante Links aufzeigen und verbessert die Zufriedenheit der Nutzer.

Es wurden in den untersuchten Studien keine konkreten Angaben zu Usability-Metriken von Automatisierten Tools gefunden. Daher wird versucht diese selbst abzuleiten. Die automatischen Tests prüfen viele Aspekte der Usability. Beispielsweise misst GTmetrix [102] die Lade-geschwindigkeit, welche Auswirkung auf die Effizienz und Zufriedenheit der Nutzer auswirkt. PowerMapper [104] testet u. a. „ob es defekte Links gibt. Dies hat Auswirkungen auf die Fehlerrate aber auch auf die Zufriedenheit der Nutzer.



- (3) *SDLC*: Klick-Analyse wird in der Betriebsphase eingesetzt. So können echte Nutzerverhalten aufgezeichnet werden. Auch wie die Klick-Analyse eignen sich Automatisierte Tools für Websites in der Betriebsphase.
- (4) *Art der Daten*: Klick Analyse kann sowohl für qualitative oder quantitative Analysen verwendet werden [32]. Qualitative Analyse deshalb, weil die Motivation der Nutzer bekannt wird, wieso die Nutzer wohin klicken. Klick Analyse generiert auch quantitative Informationen z. B. die Klick-Heatmap repräsentiert die Anzahl der Klicks.

Automatisierte Tool verführen automatische Tests durch, daher werden quantitative Informationen zur Verfügung gestellt.

- (5) *Ressourcen*: Der Recherche nach gibt es keine konkreten Angaben zum zeitlichen Aufwand bei der Klick Analyse. Doch ein Blick auf die folgende Abbildung lässt sich der Aufwand ableiten.

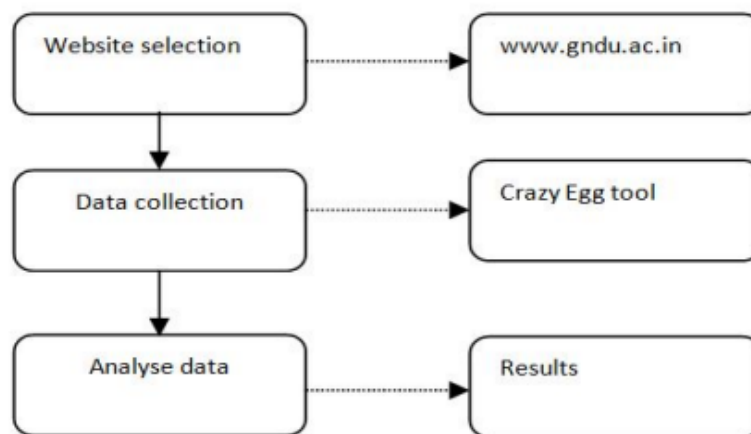


Abbildung 5.2: Klick Analyse Methodik [31]

Es sind drei Schritte bis die Ergebnisse aus einer Klick-Analyse vorliegen. Der erste Schnitt wählt eine Website aus. Im zweiten Schritt *Data Collection* werden die Daten gesammelt und das Sammeln von Daten erfolgt über Javascript, welches in den <head> oder <body> eingebunden werden kann [32]. Dieser Schritt nimmt wenig Zeit in Anspruch. Auch der dritte Schritt *Analyse data* kostet kaum Zeit weil das Tool aus den Daten beispielsweise ganz einfach eine heatmap generiert [32]. Alles in allem bin ich der Auffassung, dass der Aufwand bei der Klick Analyse gering ausfällt.

Aus eigener Erfahrung sind *automatisierte Tools* die Methode die am schnellsten und kostengünstigsten ist im Vergleich zu anderen Methoden. Es dauert weniger als einer Minute bis ein automatisierter Test mit dem entsprechenden Tools durchgeführt wird. Zudem gibt es eine Vielzahl an kostenfreien Tools, die derzeit angeboten werden wie z. B. GTmetrix Tool [102], AChecker [7], powerMapper [104], DevTool

Google Lighthouse [103] oder nibbler[106]... Im Kapitel *Tool* wird noch stärker auf die Tools eingegangen.

- (6) *Stärken*: Klick-Analyse Methode ist einfach zu verwenden und zu verstehen [32]. Des Weiteren repräsentiert Klick-Analyse das Nutzerverhalten und kann die Zufriedenheit der Nutzer verbessern [32]. Zusätzlich werden Information über die Nutzer gesammelt [56] z. B. das Land, Verweildauer oder Betriebssystem. Klick-Analyse hilft dabei herauszufinden, wie oft die einzelnen Elemente auf der Webseite angeklickt werden, ganz unabhängig davon, ob nach dem Anklicken der Elemente eine Aktion ausgelöst wird oder nicht [32]. Damit lassen sich redundante Elemente identifizieren um Maßnahmen treffen zu können. Ein weiterer Vorteil ist die Darstellung der Datenauswertung z. B. als Heatmaps, die Klick Analyse Tools anbieten. Es ist nicht besonders viel Vorwissen notwendig [25], um die Darstellung zu verstehen. Für Menschen mit kaum Wissen aus dem Bereich Web Analytics [25] können die Heatmaps trotzdem interpretieren.

Die Ergebnisse der Studie [13] zeigen, dass Automatisierte Tools eine Zeitersparnis bieten. Die große Stärke von der automatischen Analyse ist, dass die statische Code-Analyse sehr wenig Zeit in Anspruch nimmt. Nach eigenen Erfahrung beträgt die Zeit der Analyse maximal eine Minute. Danach liegt eine Zusammenfassung der Auswertung vor. Außerdem stehen eine Vielzahl an Tools zur Verfügung. Eine Menge davon sind kostenfrei, einfach zu bedienen, leicht zugänglich und bietet schnelle Ergebnisse [36].

- (6) *Schwächen*: Da es sich hier um eine statische Code-Analyse handelt, liegt keinerlei Information über das Nutzerverhalten vor. Des Weiteren wird es für jemanden ohne technischen Hintergrund (Abbildung 5.3) etwas schwierig sein, die Auswertung korrekt zu beurteilen, weil auf der Auswertung bei manchen Tools technische Hinweise stehen. Darüber hinaus haben Automatisierte Tools mindestens eine Fehlerquote von 30% [36].

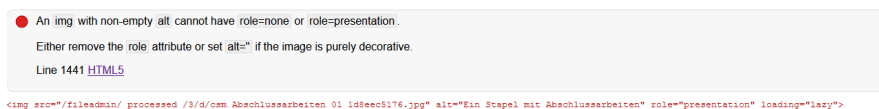


Abbildung 5.3: technischer Hinweis  
[107]

An einigen Stellen dieses Vergleichs ist kein direkter Vergleich möglich. Dafür gibt es mehrere Gründe. Zum einen haben Methoden unterschiedliche Ansätze, sodass ein direkter Vergleich nicht viel Sinn ergibt, beispielsweise Eye-Tracking mit Automatisierten Tools. Das sind zwei Methoden mit verschiedenen Herangehensweisen. Zum anderen existieren kaum Studien, in denen mehrere Methoden nicht kombiniert, sondern separat voneinander, ein interaktives Software-Produkt behandeln.

—	Klick Analyse	Automatisierte Tools
Methodentyp	Testmethode	Testmethode
USABILITY-METRIKEN	Zufriedenheit	Effizienz, Zufriedenheit, Fehler
SDLC	Betriebsphase	Betriebsphase
ART DER DATEN	qualitativ, quantitativ	quantitativ
RESSOURCEN	wenig Zeit und geringe Kosten	sehr wenig Zeit, keine bis geringe Kosten
STÄRKEN	einfach zu verwenden - verständliche Daten Visualisierung in Form von Heatmaps - gibt das Nutzerverhalten wieder - sammelt Information über alle Besucher	kostenfrei, einfach zu bedienen und leicht zugänglich, schnelle Ergebnisse
SCHWÄCHEN		Ergebnisse sind nicht 100% richtig, mindestens 30% Fehlerquote

Tabelle 5.4: Zusammenfassung des Vergleichs toolbasierter Methoden

### 5.3 ZUSAMMENFASSUNG

Für den Vergleich müssen zunächst die Vergleichskriterien als Orientierung bestimmt werden. In Bezug auf "Methoden zur Usability-Evaluation von websites" sind folgende Vergleichskriterien festgelegt worden: (1) Methodentyp, (2) Usability-Metriken (angestrebte Eigenschaften für Website-Usability), (3) Software Lebenszyklus (SDLC), (4) Art der Daten, (5) verfügbare Ressourcen, (6) Stärken und (7) Schwächen. Es wurde versucht alle Methoden an diese Vergleichskriterien zu halten. Jedoch aufgrund der unterschiedlichen Ansätze der Methoden ist es nicht immer möglich gewesen, einen typischen Vergleich durchzuführen z. B. welche Methode ist besser als die andere... Zudem wurden zu manchen Aspekten keine relevanten Ergebnisse bei der Recherche gefunden.

## SCHEMA ZUR AUSWAHL VON EVALUATIONSMETHODEN

---

Aufgrund des breiten Spektrums an Evaluationsmethoden, ist es nicht immer einfach, ein geeignetes Verfahren auszusuchen. Der Vergleich aus dem vorherigen Kapitel ist ein nützlicher Ansatz, um in der Fülle an Evaluationsmethoden zurecht zu finden. In diesem Kapitel wird versucht ein Schema zu erarbeiten um neben dem Vergleich bei der Auswahl von UEMs Abhilfe zu schaffen.

### 6.1 BESTIMMUNG DER RAHMENBEDINGUNGEN

Bevor das Schema zur Unterstützung bei der Auswahl herangezogen werden kann, ist es wichtig, die Rahmenbedingungen festzulegen. Zum einen ist es sinnvoll im Vorfeld zu klären worauf bei der Auswahl geachtet werden muss und zum anderen werden viele Fragen dabei beantwortet und können möglicherweise einige Methoden gleich zu Beginn ausgeschlossen werden. Meiner Auffassung nach werden drei Schritte nötig.

- Der erste Schritt ist den *Ist-Zustand* zu bestimmen. Angaben über den Status des Produkts sowie die Motivation sollen geklärt werden. Unterschiedliche Fragen können gestellt werden. wie z. B. Was ist das ein Produkt? In welcher Entwicklungsstufe befindet es sich? An welche Zielgruppe ist es gerichtet? Welche Probleme können aktuell bestehen? Warum soll die Usability evaluiert werden?
- Der nächste Schritt ist den *Soll-Zustand* darlegen. In diesem Schritt sollen die eventuellen Schwachstellen gezielt untersucht werden. Was soll das Produkt in Zukunft können? Welches Ziel verfolgt die Evaluation? Welche Usability-Metriken sollen verbessert werden?
- Der letzte Schritt beantworten Fragen bezüglich der *verfügbaren Ressourcen*. Wie schnell sollen die Ergebnisse geliefert werden? Wer steht zur Evaluation zur Verfügung? Wie hoch ist das Budget?

Nach der Festlegung der Rahmenbedingung stehen die Information zum Abgleichen mit dem Schema bereit.

### 6.2 METHODENAUSWAHL

Die Wahl einer Evaluationsmethode hängt von vielen Faktoren ab. Die Faktoren können mithilfe der Vergleichskriterien aus dem letzten Kapitel bestimmt werden. Der erste Faktor kann beispielsweise die aktuelle Entwicklungsstufe eines Produkts sein. Zur Vereinfachung der Darstellung (Aus-

wahl der Methode) wird der Software Entwicklungszyklus nicht detailliert in mehreren Stufen wie [82] abgebildet, sondern nur in drei Stufen: vor der Implementierung, während der Implementierung und nach der Implementierung (Betrieb) gegliedert. Zu diesen drei Stufen werden die Evaluationsklassen: expertebasierte UEM, nutzerbasierte UEM und toolbasiertes UEM zugeordnet. Expertebasierte Evaluation kann bereits in frühen Phasen der Entwicklung eingesetzt werden, wenn erste Prototypen verfügbar sind [58], Experten können viele Usability-Probleme identifizieren, bevor Nutzer die Anwendung evaluieren [58]. Experten benötigen weniger Zeit als wenn Nutzer die Anwendung testen [58]. Nutzerbasierte UEM kann auf allen Entwicklungsstufen angewendet werden, jedoch mindestens in der finalen Phase der Entwicklung [58]. Toolbasierte UEM eignet sich, wenn eine erste Version vollständig implementiert, verfügbar ist [58].

Der zweite Faktor ist, welche Art der Daten werden erhoben. Handelt es sich um quantitative oder qualitative Evaluation. Quantitative Evaluation ermöglicht eine große Anzahl an Teilnehmern [21]. Das Vorgehen erfolgt durch bestimmte Regeln [21]. Die Ergebnisse einer quantitativen Evaluation sind leicht darstellbar [21]. Qualitative Methoden dagegen sind flexibler in der Ausführung und die Ergebnisse geben ein tieferes Verständnis eines Individuums aber dafür nur eine kleine Anzahl an Teilnehmer [21]. Mehr zum Vergleich zwischen quantitativer Evaluation und qualitativer Evaluation kann in diesem Paper [21] nachgelesen werden.

Der dritte Faktor ist die Kombination aus, wer soll die Evaluation durchführen und welche Art der Daten werden erhoben. Es gibt vier mögliche Kombinationen: Experten mit qualitativ, Nutzer mit qualitativ, Nutzer mit quantitativ und Tool mit quantitativ.

Der vierte und fünfte Faktor beziehen sich auf verfügbaren Ressourcen: Zeit und materielle Kosten. Die Methoden sind untereinander relativ zu vergleichen und werden auf die Pfeile aufgetragen.

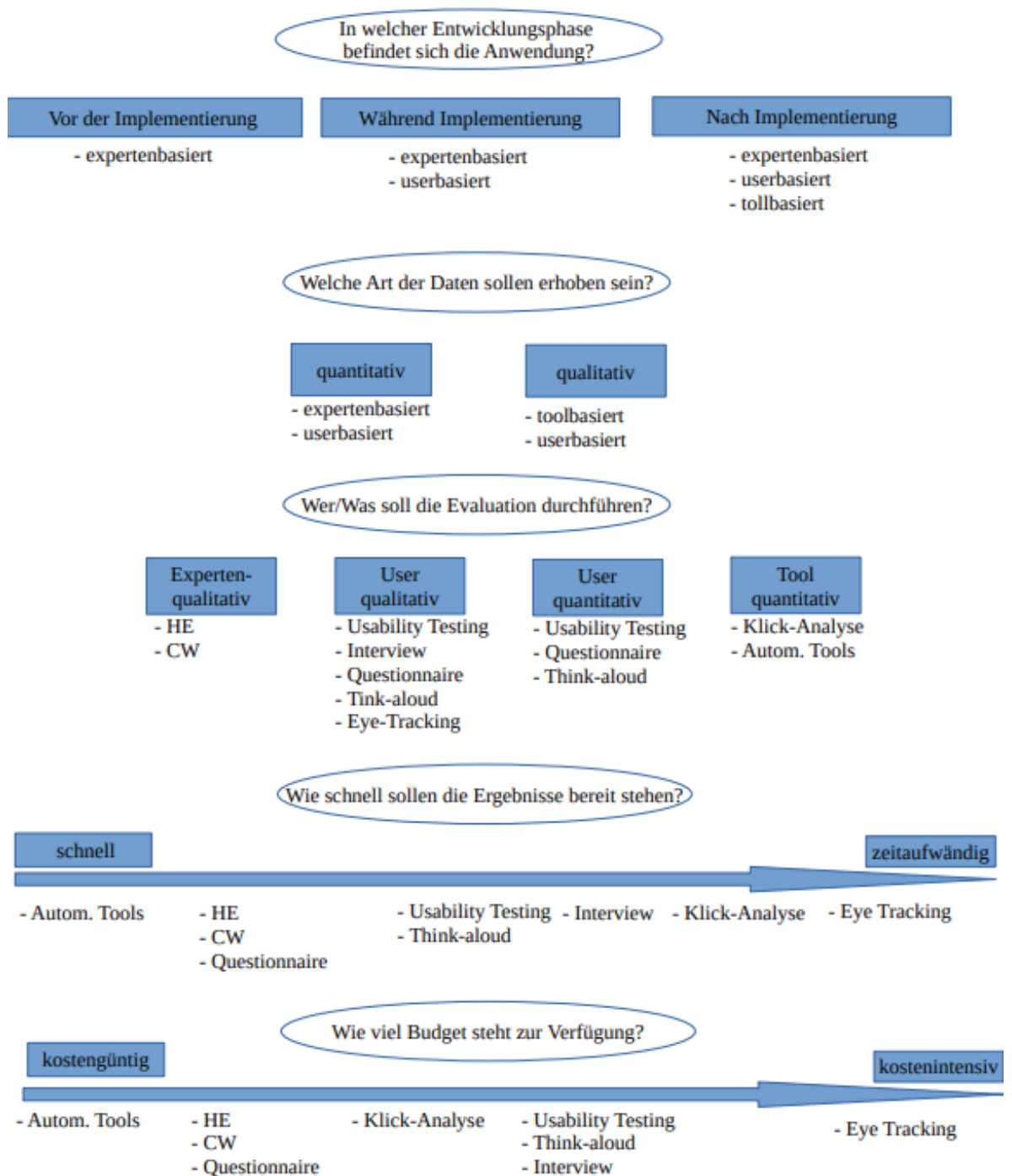


Abbildung 6.1: Schema zur Auswahl der Methode

## TOOLS

---

Nachdem die Entscheidung für eine Methode oder mehreren Methoden gefallen ist, bietet es sich an, sich über die Tools zu erkundigen, die bei dem Einsatz von Evaluationsmethoden unterstützen können. Es gibt für unterschiedliche Methoden eine Vielzahl an Anbietern. Während der Recherche verwenden Studien keine Tools für expertenbasierte Methoden. In der Regel reichen "Stift und Papier". In der Kategorie nutzerbasierte Methoden benötigt nur die Methode Eye-Tracking besondere Tools oder Geräte.

### 7.0.1 *Tools für nutzerbasierte UEM*

#### 7.0.1.1 *Eye Tracking*

Zum Setup von ET gehören u. a. Hardware (Eye-Tracker, PC sowie zwei Monitore), Software, Personal zum Durchführen von ET-Experimente [49]. Dazu ist es wünschenswert, einen speziellen und ruhigen Raum zu nehmen, um Ablenkungen oder andere Störfaktoren zu vermeiden [49].

In vielen Studien werden Eye Tracker des Anbieters Tobii verwendet [81] [87] [34] z. B. Tobii TX300, Tobii X2-60, Tobii Glasses 2. Da die Studien aus dem Jahr 2016 und 2017 stammen, sind die Modelle ausgelaufen. Derzeit bietet Tobii u. a. die neue Eyetracking-Brille *Tobii Pro Glasses 3* sowie den monitorbasierten Eyetracker *Tobii Pro Spectrum* an [92].

Bei der Verwendung der Hardware von Eye-Tracking ist auf die Beschreibung zu achten. So weist eine Studie darauf hin, dass Teilnehmer genau die Sitzhaltung entsprechend der Anleitung einnehmen [76]. Die Optimale Bedingung ist, wenn Teilnehmer still sitzen, damit die Datenerhebung so präziser wie möglich durchgeführt werden kann [76]. Mangels benötigter Ressourcen und aufgrund der umfangreichen Anforderungen in der Umsetzung ist es mir nicht möglich die Methode oder die Geräte zu testen.

### 7.0.2 *Tools für automatisierte Analyse*

#### 7.0.2.1 *Klick-Analyse*

Aus dem Vergleich der Evaluationsmethoden geht vor, dass das Betreiben von Eye Tracking mit hohen Kosten und einem hohen Zeitaufwand verbunden ist. Um aber trotzdem nicht auf die Analyse der Augenbewegungen verzichten zu müssen, schlugen Spakov und Miniotas das Verwenden von Heatmaps vor [53]. Das Aufzeichnen der Augenbewegungen durch die Klick-Analyse ist zwar nicht möglich, jedoch können mithilfe der Heatmaps von der Klick-Analyse durch das Tracken der Mausbewegungen die Blickbewegungen simuliert werden.

Eine Studie [32] hat das Tool Crazy Egg [17] zum Tracken von Mausbewegungen auf einer Website verwendet. Nach der Anmeldung auf Crazy Egg wird ein Javascript Code generiert und diesen Code kann in die Website eingebunden werden [32]. Die gesammelten Daten können in fünf Formen dargestellt werden: overlay, confetti view, scrollmap, heatmap und list [32]. Diese Darstellungen der unterschiedlichen Formen sind aus dem vorherigen Kapitel bekannt.

#### 7.0.2.2 Automatisierte Tools

In diesem Abschnitt werden einige Tools, die in den Studien verwendet wurden, vorgestellt. Diese Tools werden auf der Seite <https://fbi.h-da.de/> getestet. Um mit der Untersuchung zu beginnen, muss die URL der Website in die vorgesehene Stelle der entsprechenden Werkzeuge eingegeben werden.

- **Qualidator** [74]: Eine Studie [44] verwendet das Tool Qualidator [74] zum Messen der Usability. Die Tests basieren auf internationale Standards, Guidelines, Best Practices, ISO Normen, Usability Studien (u.a. von J. Nielsen) sowie die WCAG (Web Content Accessibility Guidelines) des World Wide Web Consortium (W3C) [72]. Das Tool analysiert die Codequalität der zu untersuchenden Website z. B. logisch gesetzte Überschriften (h1, h2, h3), keine Verwendung von Tabellen-Layout, besuchte Links werden gekennzeichnet...

Die folgende Grafik zeigt, wie das Ergebniss einer Messung aussehen kann.

### Testresultate für fbi.h-da.de

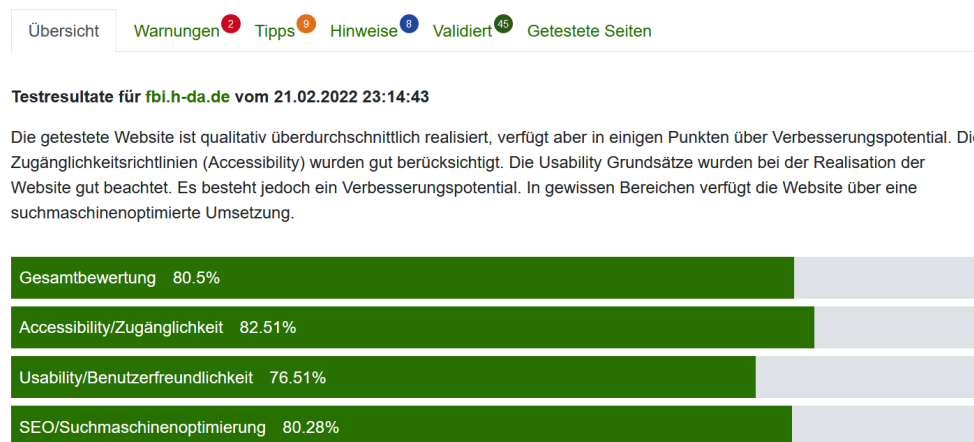


Abbildung 7.1: Resultat der Website <https://fbi.h-da.de/> durch Qualidator

Die *fbi.h-da.de* als Beispiel erzielt ein Usability-Rating von 76.51%. Das Testresultat liefert außerdem noch Tipps zur Verbesserung, Hinweise sowie ein detailliertes Ergebnis darüber, welche Tests durchgeführt wurde. Der Testbericht kann hier [73] nachgelesen werden.



Designer, Entwickler sowie Websitebetreiber können von diesem Tool profitieren [44]. Das Tool bietet eine Möglichkeit Website zu verbessern.

- *GTmetrix* [102] ist ein Usability-Testing-Tool, das die Seitengeschwindigkeit einer misst [36]. GTmetrix setzt den Schwerpunkt der Analyse auf die Performance der Webseite. GTmetrix bietet u. a. folgende Informationen:
  - *Time to Interactive*: Wie lange es dauert, bis Seite vollständig interaktiv wird.
  - *First Contentful Paint*: Wie schnell Inhalte wie Text oder Bilder auf die Seite gezeichnet werden.
  - *Largest Contentful Paint*: Wie lange es dauert, bis das größte Element auf die Seite gezeichnet ist.
  - *Speed Index*: Wie schnell die Inhalte der Seite sichtbar ausgefüllt werden.

Bis zu diesem Zeitpunkt (03.02.2022) hat das Tool GTmetrix bereits die Performance bei rund 872,7 Millionen Webseiten gemessen. Die Seite des Fachbereichs Informatik erzielt laut des Tools eine Performance von ca. 76%, entspricht einer Note C im amerikanischen Notensystem von A bis F, wobei A die beste Note ist. Auch dieses Tool gibt Hinweise, an welcher Stelle der Website nicht gut geschnitten hat und bietet Verbesserungsvorschläge an.

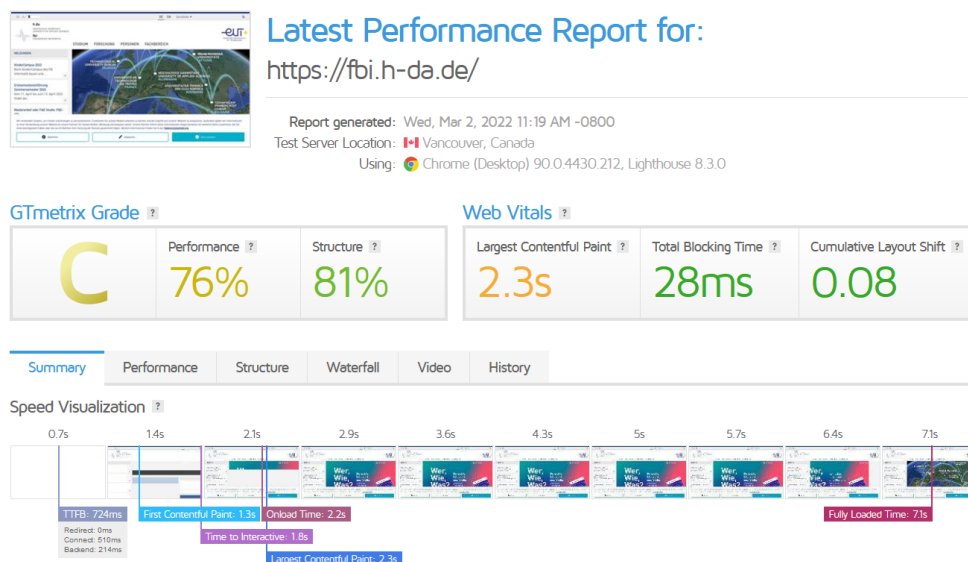


Abbildung 7.2: Resultat der Website <https://fbi.h-da.de/> durch GTmetrix

- *nibbler*[106]: Dieses Tool ist übersichtlich. Auf der rechten Seite der Auswertung stehen die Punkte, die das Tool getestet hat. Beispielsweise prüft das Tool:
  - ob die Seite responsive ist,

- ob die Images nach den neusten Best Practice eingefügt wurden,
- ob die Seite Analytics verwendet,
- oder ob die Seite für den Druck geeignet ist.

Die Tests des Tools sind noch nicht umfangreich. Außerdem gibt das Tool keine Auskunft darüber, auf welche Guidelines oder Best Practices die Tests basieren. Das Tool verweist aber darauf hin, dass die Ergebnisse der Analyse nicht korrekt sein können und den Erwartungen nicht unbedingt entsprechen [105].

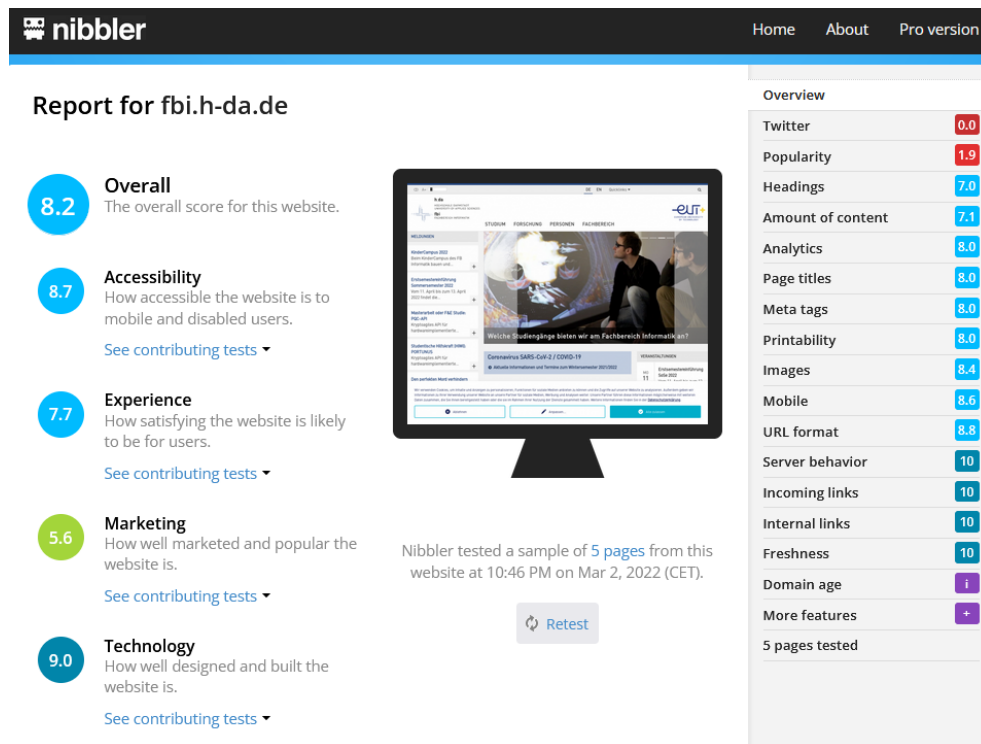


Abbildung 7.3: Resultat der Website <https://fbi.h-da.de/> durch Nibbler

- *PowerMapper*[104]: Anders als Nibbler basieren die Tests auf die Guidelines *Usability.gov* [93] und *W3C Best Practices*. Auch wie andere Tools wie Qualidator, Nibber, GTmetrix verweist PowerMapper auch auf die Fehler und gibt Hilfestellung zum Verbessern. Die Abbildung 7.4 ist die Auswertung der Seite *fbi.h-da.de* von PowerMapper.

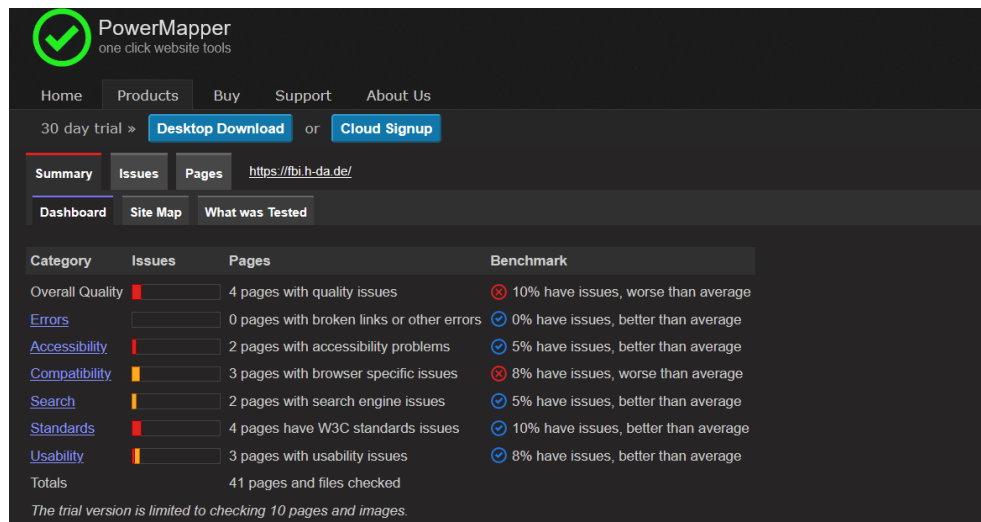


Abbildung 7.4: Resultat der Website <https://fbi.h-da.de/> durch PowerMapper

Weitere Tools sind: Pingdom zur Messung von Ladezeit und Seitengröße, Websitepulse für defekte Links, Webpagefx Tool für Lesbarkeit [26], Lighthouse ein Devtool von Google.

Die automatische Toolsanalyse deckt u. a. die Kernpunkte der Usability. Jedoch ist es zu beachten, dass dieser Test keinesfalls manuelle Tests oder Inspektion ersetzt, da diese Tests mit einer Fehlerquote von mindestens 30% zu rechnen ist [36].

## SCHLUSS

---

### 8.0.1 Diskussion

Die Recherche hat 67 Studien untersucht mit dem Ziel herauszufinden, welche Usability-Evaluationsmethoden derzeit am meisten verwendet werden. Die Untersuchung kam zu dem Ergebnisse: Questionnaire, Heuristische Evaluation und Automatisierte Tools sind die drei am häufigsten verwendeten Verfahren. Da stellt sich die Frage, warum diese drei Methoden so beliebt sind.

Mit Hilfe des Vergleichs wird versucht die Frage zu beantworten. Zunächst fällt auf, dass bei die Ansätze der drei Methoden *leicht zu verstehen* sind. Die Methoden sind unkompliziert in jeder Hinsicht. Die Evaluation mit Questionnaire erfolgt durch eine Liste von Fragen. Heuristische Evaluationen benötigen nur die Heuristiken und drei bis fünf Experten und Automatisierte Tools brauchen nur eine URL für die automatischen Tests.

Des Weiteren funktionieren die Methoden nach einem *festgelegten Konzept*. Bei Questionnaire gibt es vorbereitete Fragen und ein eigenes Berechnungssystem. So können die Anwender gleich loslegen ohne mühsame Vorbereitung. Auch der HE stehen die Heuristiken (Prinzipien) zur Verfügung z. B. die Heuristiken von Nielsen. Je nach Anwendungsfall können die Heuristiken bei Bedarf angepasst oder erweitert werden. Bei Automatisierten Tools ist ebenfalls keine Vorbereitung notwendig, da die Tools nur die URL benötigen und die Tests werden automatisch durchgeführt.

Darüber hinaus werden bei den drei Methoden, in Hinblick auf die anderen Methoden aus dem Vergleich, *mehr Usability-Metriken* bei der Evaluation berücksichtigt. So decken Questionnaire und Heuristische Evaluation die meisten Metriken ab.

Weiter beanspruchen die Methoden *wenig Ressourcen*. Die Methoden benötigen wenig Zeit und können schnell Ergebnisse liefern.

Doch diese Methoden haben auch Einschränkungen. In den Studien ist eines festzustellen, Questionnaire und Automatisierte Tools werden nicht als eigenständige Methoden zum Evaluieren benutzt. Bei Questionnaire fehlen die praktischen Testdurchführungen und bei Automatisierten Tools liegt die Fehlerquote bei mindestens 30% [36]. Auch manche Tools weisen darauf hin, dass die automatische Analyse die manuellen Tests nicht ersetzen können. Jedoch können Questionnaire und automatisierte Tests anderen Methoden als Ergänzung dienen.

Im Vergleich zu anderen Kategorien hat die Kategorie *nutzerbasierte UEM* mehr Methoden. Meiner Meinung nach liegt es daran, weil die Nutzertests unerlässlich sind. Schließlich werden die interaktiven Software Produkte für die Endnutzer gemacht. Außerdem werden bei nutzerbasierte UEM ande-

re Usability-Probleme als beispielsweise bei expertenbasierten UEM [75]. Es gibt zwar Überschneidungen aber der Anteil an einmaligen Usability-Probleme ist höher als der Anteil bei der Überschneidung. Es hat wahrscheinlich einen einfachen Grund, weil es zwei unterschiedliche Gruppen sind und jede Gruppe guckt mit einem anderen Blickwinkel auf die Anwendung.

Das Eye-Tracking in dieser Kategorie benötigt im Vergleich zu anderen Methoden nicht nur eine besondere Software aber auch besondere Geräte. Das Setup von ET ist anders als bei anderen Verfahren umfangreich. Zudem setzt ET mindestens 46 Teilnehmer voraus um aussagekräftig zu sein. Dadurch ist ET sehr aufwändig und beansprucht viel Ressourcen.

Eine Alternative für ET im Bereich Website kann meiner Meinung nach die Klick-Analyse sein. Das Tool für die Klick-Analyse ist viel günstiger oder gar kostenfrei z. B. hotjar [35] und stellt auch Heatmaps zur Verfügung. Da unsere Augen beim Bedienen einer Website die Mausbewegungen verfolgen, funktioniert das Maus-Tracking fast analog wie Eye-Tracking. Zwar stimmen unsere Augenbewegungen nicht mit den Mausbewegungen vollständig überein, wenn die Maus stehen bleibt und unsere Augen sich weiterhin über die Website bewegen. Aber meiner Einschätzung nach ist die Klick Analyse eine gute Alternative um ET zu ersetzen, wenn ET aus Ressourcengründen nicht möglich ist.

In den 67 untersuchten Studien wurden mindestens 89 Methoden festgestellt. Es lässt sich darauf schließen, dass in zahlreichen Studien mehr als eine Methode zum Einsatz kam. Jede Methode hat seine Stärken und Schwächen, und jede Methode hat unterschiedliche Ansätze zum Evaluieren von Websites. Daher ist es zu empfehlen, von Fall zu Fall zu prüfen, ob mehrere Methoden infrage kommen. Es ist in vielen Fällen von Vorteil die Methoden zu kombinieren um umfassende, validere, verlässlichere und objektivere Daten zu erhalten [27], [39].

### 8.0.2 Zusammenfassung

Im Rahmen dieser Arbeit sollten aktuelle Usability-Evaluationsmethoden untersucht werden, die derzeit verwendet werden. Die Methoden sind anhand von verschiedenen Kriterien miteinander zu vergleichen mit dem Ziel für einen bestimmten Anwendungsfall eine geeignete Methode zu finden.

Zu Beginn der Arbeit wurden 67 Studien von 2015-2021 untersucht um die aktuellen Verfahren zu identifizieren. Dabei wurden eine Vielzahl an Methoden festgestellt. Die drei häufigsten Methoden sind: Questionnaire, Heuristische Evaluation und Automatisierte Tools.

Aufgrund des breiten Spektrum an Evaluationsmethoden wurden sie in Kategorien klassifiziert um sie zu strukturieren. Dabei sind drei Gruppen entstanden: expertenbasierte Usability-Evaluationsmethoden (UEMs), nutzerbasierte UEMs und toolbasierte UEMs. Danach wurden die Methoden in die entsprechende Kategorie zugeordnet und vorgestellt.

Im nächsten Schritt wurden die ausgewählten Methoden aus der Recherche miteinander verglichen. Um dies zu erreichen wurden Vergleichskriteri-

en rund um das Thema Evaluationsmethoden zum Evaluieren von Website-Usability festgelegt. Dabei stellte sich heraus, dass aufgrund der unterschiedlichen Ansätze der Methoden der Vergleich nicht an allen Punkten möglich war.

Auf Basis dieser Gegenüberstellung der Methoden wurde ein Schema zur Methodenauswahl erstellt. Vor der Nutzung dieses Schemas sollten Anwender einige Fragen vorher beantworten um das Schema besser nutzen zu können. Mit diesem Schema ist es möglich eine geeignete Methode für seinen Anwendungsfall aus einer Vielzahl von Methoden systematisch herauszufinden.

Zum Schluss werden einige Tools vorgestellt, die gemeinsam mit den Methoden beim Evaluieren von Website-Usability helfen sollen.

Diese Arbeit bietet Unterstützung für Anwender, die ihre Website-Usability verbessern wollen und vor der Entscheidung stehen, welche Methoden existieren und welche derzeit angeboten oder benutzt werden und wie lässt sich eine geeignete Methode aus der Menge auswählen.

### 8.0.3 *Ausblick*

Die Arbeit hat gezeigt, dass derzeit eine Vielzahl von Evaluationsmethoden angeboten wird. Um eine geeignete Methode zu finden, sollte ein Überblick über das Thema geschaffen werden. Des weiteren hat die Arbeit auch gezeigt, dass eine Methode nicht ausreichen wird, um die Usability einer Website zu verbessern sondern es von Vorteil ist, mehrere Methoden für die Evaluation miteinander zu kombinieren um das bestmögliche Ergebnis zu erzielen.

Diese Arbeit hat für die Recherche nach aktuellen Tools nur Studien, bei denen die Usability von Websites werden, untersucht. Es stellt sich die Fragen, ob diese Methoden auch für andere interaktive Software verwendet werden können.

Des Weiteren stellte sich heraus, dass es bewährte Methoden gibt, die in den 1990er Jahren entwickelt wurden und derzeit immer noch sehr stark Anwendung finden. Diese werden auch in Zukunft relevant bleiben. Das sind beispielsweise die Heuristische Evaluation und die verschiedenen Typen von Questionnaires. Darüber hinaus ist eines zu erkennen, immer mehr Studien nutzen die Automatisierten Tools als Ergänzung zu anderen traditionellen Methoden um die Ergebnisse zu validieren. In der Zukunft ist es denkbar, dass die Automatisierten Tools sichere Ergebnisse liefern und die Fehlerquote der Analyse reduzieren, damit Anwender stärker auf die automatischen Tests sich verlassen können.

## LITERATUR

---

- [1] H. Fawareh und A. Al-Bardeen. *An Investigation Of E-learning Websites Usability Issues during the COVID-19 Pandemic: Jordan Case Study*. IEEE, 2021.
- [2] J. Sheikh und A. Arshad. *Using Heuristic Evaluation to Enhance the Usability: A Model for Illiterate Farmers in Pakistan*. Springer International Publishing AG, 2018.
- [3] N. Handayani W Kusuma Z. Rosyada Y. Widharto und A. Hanifah. *Usability Evaluation of Inventory Information System Design of Disaster Management in Yogyakarta Province - Indonesia*. Association for Computing Machinery, 2020.
- [4] S. Ritthiron und A. Jiamsanguanwaong. *Usability Evaluation of the University Library Network's Website Using an Eye-Tracking Device*. ACM, 2017.
- [5] A. Doubleday M. Ryan M. Springett und A. Sutcliffe. *A comparison of usability techniques for evaluating design*. DIS'97, 1997.
- [6] Y. Septiandi und A. Suzianti. *User Interface Redesign of Technology Based Learning Service using Usability Testing Methods. Case Study: Rumah Belajar*. Association for Computing Machinery, 2019.
- [7] *AChecker Tool*. URL: <https://achecker.achecks.ca/checker/index.php> (besucht am 27.02.2021).
- [8] Anas Abulfaraj und Adam Steele. *Operational Usability Heuristics: A Question-Based Approach for Facilitating the Detection of Usability Problems*. Springer Nature Switzerland AG, 2021.
- [9] I. Abuqaddom H. Alazzam A. Hudaib F. Al-Zaghoul. *A measurable website usability model: Case Study University of Jordan*. IEEE, 2019.
- [10] Martin Armstrong. *How Many Websites Are There?* 6. Aug. 2021. URL: <https://www.statista.com/chart/19058/number-of-websites-online/> (besucht am 09.01.2021).
- [11] M. Blackmon P. Polson M. Kitajima und C. Lewis. *Cognitive Walkthrough for the Web*. Association for Computing Machinery, Minneapolis, Minnesota, USA, 2002.
- [12] F. Paz F. A. Paz J. Arenas und C. Rosas. *A Perception Study of a New Set of Usability Heuristics for Transactional Web Sites*. Springer International Publishing AG, 2018.
- [13] N. Ismail F. Jamaluddin A. Hamidan A. Ali S. Mohamed und C. Said. *Usability Evaluation of Encyclopedia Websites*. UTM, International Journal of Innovative Computing, 2021.

- [14] Elizabeth Charters. *The Use of Think-aloud in Qualitative Research An Introduction to Think-aloud Methods*. Brock Education, 2016.
- [15] Nurul R Nurwulan Chesia D Sara. *Comparative Usability Evaluation of Novice and Expert Gojek Users*. Association for Computing Machinery, 2021.
- [16] D. Lavery G. Cockton und M. Atkinson. *Comparison of evaluation methods using structured usability problems reports*, *Behavior & Information Technology*, 16:4-5, 246-266, DOI: 10.1080/014492997119824. Taylor & Francis, 1997.
- [17] Crazy Egg. URL: <https://www.crazyegg.com/> (besucht am 16. 02. 2021).
- [18] D. Hix J. Gabbard J. Swan M. Livingston T. Höllner S. Julier Y. Baillot und D. Brown. *A Cost-Effective Usability Evaluation Progression for Novel Interactive System*. In *Proceedings Hawaii International Conference on Systems Sciences*, 2004.
- [19] Harmut Ditton. *Evaluation und Qualitätssicherung*. Springer Fachmedien Wiesbaden, 2002.
- [20] P Sukmasetya A. Setiawan und E. Armi. *Usability evaluation of a university website: a case study*. IOP Publishing Ltd, 2020.
- [21] F. Butterfoss V. Francisco und E. Capwell. *Choosing Effective Methods*. Sage Publications, Inc., 2000.
- [22] T. Urgas S. Gülsecen C.Cubukcu I. Erdohmus. *Research Trends in Web Site Usability: A Systematic Review*, S.517-528. Springer International Publishing Switzerland, 2016.
- [23] *Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability*. 1998.
- [24] R. Pradini R. Kriswibowo und F. Ramdani. *Usability Evaluation on the SIPR Website Uses the System Usability Scale and Net Promoter Score*. IEEE, 2021.
- [25] T. Farney. *Click Analytics: Visualizing Website Use Data*. 2010.
- [26] M. Singh und D. Dani G. Agrawal D. Kumar. *Evaluating Accessibility and Usability of Airline Websites*. Springer Nature Singapore Pte Ltd., 2019.
- [27] P. Dewi G. Dantes und G. Indrawan. *User experience evaluation of e-report application using cognitive walkthrough (cw), heuristic evaluation (he) and user experience questionnaire (ueg)*. *Journal of Physics: Conference Series*, 1516 012024, 2020.
- [28] C. Ohl und G. Schade. *Klassifikationsschema für Usability-Evaluationsmethoden*. Diefenbach, Henze, Pielot, 2015.
- [29] Sugandha Gupta. *A Comparative study of Usability Methods*. *International Journal of Computer Trends und Technologys*, 2015.



- [30] R. Fukuda und H. Bubb. *Eye tracking study on Web-use: Comparison between younger and elderly users in case of search task with electronic timetable service*. PsychNology Journal, 2003.
- [31] K. Kaur und H. Singh. *Analysis of Website using Click Analytics*. IJCSET, 2015.
- [32] K. Kaur und H. Singh. *Click Analytics: What clicks on Webpage indicates?* 2nd International Conference on Next Generation Computing Technologies, 2016.
- [33] H. Hamarashid und S. Saeed. *Usability Testing on Sulaimani Polytechnic University Website*. International Journal of Multidisciplinary und Current Research, 2017.
- [34] O. Hareide und R. Ostnes. *Maritime Usability Study by Analysing Eye Tracking Data*. They Royal Institute of Navigation, 2017.
- [35] *Heatmap Tool: Hotjar*. URL: <https://www.hotjar.com/> (besucht am 03.03.2022).
- [36] B. Csontos und I. Heckl. *Accessibility, usability, and security evaluation of Hungarian government websites*. SpringerLink, 2020.
- [37] M. Azhar E. Krisnanik R. Wirawan und I. Indriana. *UI Analysis and Redesign of A Simpul Seribu Website Using Cognitive Walkthrough Method*. IEEE, 2021.
- [38] P. Manakhov V. Ivanov. *Defining usability Problems*. P. Manakhov V. Ivanov, 2016.
- [39] J. Kontio L. Lehtola und J. Bragge. *Using the Focus Group Method in Software Engineering: Obtaining Practitioner and User Experiences*. IEEE, 2004.
- [40] F. Paz F.A. Paz und J. Poe-Sang. *Experimental Case Study of New Usability Heuristics*. Springer International Publishing Switzerland, 2015.
- [41] Rolf Molich Jakob Nielsen. *10 Usability Heuristics for User Interface Design*. 15. Nov. 2021. URL: <https://www.nngroup.com/articles/ten-usability-heuristics/> (besucht am 17.12.2021).
- [42] Freddy Paz und Jose Antonio Pow-Sang. *Current Trends in Usability Evaluation Methods: A Systematic Review*. IEEE, 2014.
- [43] N. Aziz und K. Adzhar. *Development of Website Usability Instrument Based on Expert Review*. Aziz2015DevelopmentOW, 2015.
- [44] S. Kaur K. Kaur und P. Jusoh. *Analysis of Website Usability Evaluation Methods*. IEEE, 2016.
- [45] N Marwah und L-Kharb. *EYE TRACKING: A NEW EVOLUTION IN TECHNOLOGY*. International Journal of Engineering Applied Sciences und Technology, 2019.
- [46] I. Beck L. Kucan. *Thinking Aloud and Reading Comprehension Research: Inquiry, Introduction, and Social Interaction*. Sage Journals, 1997.

- [47] T. Saba und a. Mateen M. Ashraf F. Cheema. *Usability of Government Websites*. (IJACSA) International Journal of Advanced Computer Science und Applications, 2017.
- [48] A. Namoun und A. Taleb M. Benaida. *Evaluation of the Impact of Usability in Arabic University Websites: Comparison between Saudi Arabia and the UK*. (IJACSA) International Journal of Advanced Computer Science und Applications, 2018.
- [49] Z. Sharafi B. Sharif Y. Gueheneuc A. Begel R. Bednarik und M. Crosby. *A practical guide on conducting eye tracking studies on software engineering*. Springer Science Business Media, LLC, part of Springer Nature, 2020.
- [50] C. Chung M. Fan J. Lin und K. Truong. *Concurrent Think-Aloud Verbalizations and Usability Problems*, ACM Transactions on Computer-Human Interaction, Vol. 26, No. 5, Article 28. Association for Computing Machinery, 2019.
- [51] H. Tadayon und M. Jabali M. Farzandipour E. Nabovati. *Comparison Usability Evaluation Methods in a Health Information System: Heuristic Evaluation versus Cognitive Walkthrough Method*. Research Square, 2021.
- [52] A.H. Kusumo und M. Hartono. *The evaluation of academic website using eye tracker and UEQ: a case study in a website of xyz*. IOP Conference Series: Materials Science und Engineering, 2019.
- [53] I. Kirsh und M. Joy. *A Different Analytics Perspective Through Copy Clipboard Heatmaps*. Springer Nature Switzerland AG, 2020.
- [54] U. Frick M. Tallon M. Reichert M. Winter H. Baumeister und R. Pryss. *Exploring the Usability of the German COVID-19 Contact Tracing App in a Combined Eye Tracking and Retrospective Think Aloud Study*. IEEE, 2021.
- [55] L.W. Peute S.L. Knijnenburg L.C. Kremer und M.W.M. Jaspers. *A Concise and Practical Framework for the Development and Usability Evaluation of Patient Information Websites*. Schattauer, 2015.
- [56] Sadi E. Nursubiyantoro Y. Astanti Ismianti A. Wibowo H. Mastrisiswadi. *Usability Evaluation of Scientific Journal Websites using the System Usability Scale (Case Study of the OPSI Journal Website)*. Research Synergy Foundation, 2021.
- [57] Raphael Philipp Menges. *Improving Usability and Accessibility of the Web with Eye Tracking*. Universität Koblenz Landau, WeST People und Knowledge Networks, 2021.
- [58] H. Petrie und N. Bevan. *The Evaluation of Accessibility, Usability, and User Experience*. The Universal Access Handbook, C Stepanidis (ed), CRC Press, 2009.

- [59] A. Shekhar und N. Marsden. *Cognitive Walkthrough of a Learning Management System with Gendered Personas*. Association for Computing Machinery, New York, NY, United States, 2018.
- [60] H. Tüzün und M. Yildiz N. Menzi-Cetin E. Alemdag. *Evaluation of a university website's usability for visually impaired students*. Springer-Verlag Berlin Heidelberg, 2015.
- [61] M. Lakshman L. Sinha M. Biswas M. Charles und N.K. Arora. *Quantitative Vs Qualitative Research Methods*. The Indian Journal of Pediatrics, 2000.
- [62] J. Nielsen. *Usability Inspection Methods*. Association for Computing Machinery, 1994.
- [63] J. Nielsen. *Usability Metrics: Tracking Interface Improvements*. IEEE, 1996.
- [64] Jakob Nielsen. *Usability Heuristics for User Interface Design*. 15. Nov. 2020. URL: <https://www.nngroup.com/articles/ten-usability-heuristics/> (besucht am 03.02.2022).
- [65] S. Khodambashi und Oystein Nytro. *Usability Evaluation of Puplished Clinical guidelines on the Web: A Case Study*. IEEE, 2016.
- [66] O. Ahhadreti und P. Mayhew. *Are Two Pairs of Eyes Better Than One? A Comparison of Concurrent Think-Aloud and Co-Participation Methods in Usability Testing*. JUS, Journal of Usability Studies, 2018.
- [67] O. Alhadreti und P. Mayhew. *To Intervene or Not to Intervene: An Investigation of Three TA Protocols in Usability Testing*. 2017.
- [68] O. Alhadreti und P. Mayhew. *Rethinking Thinking Aloud: A Comparison of Three Think-Aloud Protocols*. ACM, Montreal, Canada, 2018.
- [69] A. Setiawan P. Sukmasetya E. Arumi. *Combining Usability Testing with In-Depth Interview for Online Credit Hour Website Evaluation*. khazanah informatika, 2020.
- [70] Kara Pernice. *Text Scanning Patterns: Eyetracking Evidence*. 25. Sep. 2019. URL: <https://www.nngroup.com/articles/text-scanning-patterns-eyetracking/> (besucht am 04.01.2021).
- [71] Witold Suryn Prof. Motoei AZUMA Nigel BEVAN. *ISO/IEC 9126-4 Software Engineering - Software product quality*. Version 6.0. International Standard, 2001.
- [72] Qualidator About. URL: <https://www.qualidator.com/de/ueber-uns/der-websitequalitymonitor.htm> (besucht am 21.02.2022).
- [73] Qualidator Tool Testergebnis fbi.h-da.de. URL: <https://www.qualidator.com/de/myrating/de/h/-/d/fbi.h-da.de/index.htm> (besucht am 21.02.2022).
- [74] Qualidator Tool. URL: <https://www.qualidator.com/de/tools/myrating-overview.htm> (besucht am 21.02.2022).
- [75] W. Tan D. Lui und R. Bishu. *Web evaluation: Heuristic evaluation vs user testing*. International Journal of Industrial Ergonomics, 2009.

- [76] D. Niehorster T. Cornelissen K. Holmqvist I. Hooge und R. Hessels. *What to expect from your remote eye-tracker when participants are unstrained*. Springerlink, 2017.
- [77] J. Nielsen und R. Molich. *Heuristic evaluation of user interfaces*. In: CHI '90: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Washington, Seattle, USA, 1990.
- [78] I. Sikiru und A. Balogun R. Oyekunle O. Bello Q. Jubril. *Usability Evaluation using Eye-Tracking on E-Commerce and Education Domains*. Journal of Information Technology und Computing, 2020.
- [79] M. Winter H. Baumeister U. Frick M. Tallon M. Reichert und R. Pryss. *Exploring the Usability of the German COVID-19 Contact Tracing App in a Combined Eye Tracking and Retrospective Think Aloud Study*. International Conference of the IEEE Engineering in Medicine Biology Society (EMBC), 2021.
- [80] F. Firjatullah J. T. Haryoko und R. Zuraida. *Remote Usability Testing Evaluation on the Most Visited E-commerce Website in Indonesia During Covid-19 Pandemic*. IEEE, 2021.
- [81] P. Weichbroth K. Redlarski und I. Garnik. *Eye-tracking Web Usability Research*. IEEE, 2016.
- [82] N. Ruparelia. *Software Development Lifecycle Models*. Association for Computing Machinery, New York, NY, United States, 2010.
- [83] E. Elberkawi N. El firjani A. Maatuk und S. Aljawarneh. *Usability Evaluation of Web-based Systems: A New Method and Results*. IEEE, 2016.
- [84] H. Gull und S. Iqbal. *Usability Evaluation of E-Government Websites in Saudi Arabia by Cognitive Walkthrough*. IGI Global, 2017.
- [85] T. Ramayah Sardar Zafar Iqbal Saqib Saeed Yasser A. Bamarouf. *The State of Art in Website Usability Evaluation Methods*, S.275-296. IGI Global, 2017.
- [86] Arslan Sheikh. *Evaluating the usability of COMSATS Institute of Information Technology library website*. Emerald Publishing Limited, 2017.
- [87] Julia Falkowska Janusz Sobiecki und Martyna Pietrzak. *Eye Tracking Usability Testing Enhanced with EEG Analysis*, S.399-411. Springer International Publishing Switzerland, 2016.
- [88] Jan M. Noyes und Stella Mills. *Heuristic Evaluation of Websites: Lessons Learned*. International Journal of The Computer, the Internet und Management Vol. 14.No.1, 2006.
- [89] Suchmaschine Semantic Scholar. URL: <https://www.semanticscholar.org/> (besucht am 26.02.2021).
- [90] B. Soriano S. Duarte A. Alonso A. Perales und T. Carreno. *Users evaluation of a Spanish eHealth pediatric website Survey Study*. JMIR Publications, The Leading eHealth Publisher, 2020.

- [91] S. Habibi L. Seyed-Akbari A. Torab-Miandoab und T. Samad-Soltani. *Usability of Central Library Websites of Iranian Universities of Medical Sciences: An Evaluation*. DESIDOC Journal of Library Information Technology, 2019.
- [92] Tobii Glasses 3. URL: <https://www.tobiipro.com/de/produkte/hardware/> (besucht am 03.03.2021).
- [93] Usability Guideline: Usability.gov. URL: <https://www.powermapper.com/products/sortsite/rules/usegov5.1/> (besucht am 03.03.2022).
- [94] A. Baravalle und V. Lanfranchi. *Remote Web usability testing*. Psychonomic Society, Inc., 2003.
- [95] C. Rusu V. Rusu S. Roncagliolo J. Apablaza und V.Z. Rusu. *User Experience Evaluation: Challenges for Newcomers*. Springer International Publishing Switzerland, 2015.
- [96] N. Chuan A. Sivaji und W. Ahmad. *Usability Heuristics for Heuristic Evaluation of Gestural Interaction in HCI*. Springer International Publishing Switzerland, 2015.
- [97] E. Wang und B. Caldwell. *An Emperical Study Of Usability Testing: Heuristic Evaluation vs. User Testing*. Proceedings of the Human Factors und Ergonomics Society Annual Meeting, 2002.
- [98] Michael Witzenleiter. *Wie Sie Ihr Website- und E-Commerce-Testing erfolgreich auf- und umsetzen*. Springer Gabler, Wiesbaden, 2021.
- [99] World Wide Web Consortium. URL: <https://www.w3.org/> (besucht am 03.02.2022).
- [100] N. Aziz N. Sulaiman W. Hassan N. Zakaria A. Yaacob. *A Review of Website Measurement for Website Usability Evaluation*. J. Phys.: Conf. Ser. 1874 012045, 2021.
- [101] S. McDonalds H. Edwards T. Zhao. *Exploring Think -Aloud in Usability Testing: An International Survey*. IEEE Transactions on Professional Communication, 2012.
- [102] automatische Analyse-Tool GTmetrix. URL: <https://gtmetrix.com/> (besucht am 03.02.2022).
- [103] automatische Analyse-Tool Google Lighthouse. URL: <https://developers.google.com/web/tools/lighthouse> (besucht am 03.02.2022).
- [104] automatische Analyse-Tool PowerMapper. URL: <https://www.powermapper.com/> (besucht am 03.02.2022).
- [105] automatische Analyse-Tool nibbler. URL: [https://nibbler.silktide.com/en\\_US/about](https://nibbler.silktide.com/en_US/about) (besucht am 03.02.2022).
- [106] automatische Analyse-Tool. URL: [https://nibbler.silktide.com/en\\_US](https://nibbler.silktide.com/en_US) (besucht am 03.02.2022).
- [107] hotjar.com. *Types of Heatmaps*. URL: <https://help.hotjar.com/hc/en-us/articles/115011867048-Types-of-Heatmaps> (besucht am 11.01.2021).

- [108] J. Lamberz T. Litfin G. Meeh bunse und Ö. Teckert. *User-Friendly Website Design A Combined Eye-Tracking Study*. Dubrovnik, Croatia, 2017.