



Universität Regensburg

**Evaluation im Reallabor: Analyse und
Anwendung von Evaluationsmethoden der
Mensch-Maschine-Interaktion in Reallaboren im
Rahmen des Forschungsprojekts
„Strukturvorgaben für den Parteivortrag“**

Masterarbeit im Fach Medieninformatik
am Institut für Information und Medien, Sprache und Kultur (I:IMSK)

Vorgelegt von:	Marie Sautmann
Adresse:	████████████████████
E-Mail (Universität):	marie.sautmann@stud.uni-regensburg.de
E-Mail (privat):	marie.sautmann@gmx.de
Matrikelnummer:	████████
Erstgutachter:	Prof. Dr. Christian Wolff
Zweitgutachter:	Dr. Raphael Wimmer
Betreuerin:	Julia Sageder, M.Sc.
Laufendes Semester:	7. Semester M.Sc. Medieninformatik, 1. Semester M.A. Ver- gleichende Kulturwissenschaft
Abgegeben am:	31.12.2023

Inhaltsverzeichnis

Aufgabenstellung	7
1. Einleitung	8
2. Projektverlauf und Forschungsauftrag „Strukturvorgaben für den Parteivortrag“	12
3. Die Methode Reallabor	19
3.1. Reallabore international	20
3.2. Reallabore in Deutschland	22
3.2.1. Entstehung und Definition	22
3.2.2. Reallaborlandschaft in Deutschland: Ein Überblick	24
3.3. [Exkurs] Digitalisierung in der (bayrischen) Justiz	28
3.4. Reallabor „Strukturvorgaben für den Parteivortrag“	30
4. Wahrnehmung des Projekts Basisdokument	33
4.1. Wahrnehmung des Basisdokuments in der juristischen Fachwelt	33
4.2. Wahrnehmung der Erprobung im Reallabor Strukturvorgaben	37
4.3. Wahrnehmung des Basisdokumentprototyps	38
5. Projektstand im August 2023	41
5.1. Methode: Fokusgruppe mit Projektbeteiligten	41
5.2. Ergebnisse der Fokusgruppe	42
5.2.1. Projektziele	43
5.2.2. Aktuelle Probleme im Reallabor	43
5.2.3. Erwartungen an die Evaluation	44
5.3. Neuere Entwicklungen im Projekt (Dezember 2023)	45
6. Forschungsmethoden der Mensch-Maschine-Interaktion	47
6.1. User Centered Design	48
6.2. Kontrollierte Experimente	48
6.3. Umfragen	49
6.4. Tagebuchstudien	50
6.5. Fallstudien	50
6.6. Interviews und Fokusgruppen	51
6.7. Ethnographie	52
6.8. Usability Testing	52
6.8.1. Methoden des expert:innenbasierten Testens	53
6.8.2. Automatisiertes Testen	54
6.8.3. Methoden des nutzer:innenbasierten Testens	54
7. Evaluation im Feld	57
7.1. Feldversuche in der HCI	58
7.2. Evaluationsmethodik in der Feldforschung	61

8. Evaluation im Reallabor Strukturvorgaben: Planung der Projektgruppe	62
9. Anwendung von Evaluationsmethoden auf das Reallabor Strukturvorgaben	65
9.1. Kriterien- und Fragenkatalog	67
9.2. Ergebnisse der Analyse	70
9.3. Diskussion der Analyseergebnisse	71
10. Diskussion und Fazit	77
11. Ausblick	82
Literaturverzeichnis	85
A. Anhang	98
A.1. Beispiele für standardisierte Fragebögen	98
A.2. Fragenkatalog für die Evaluation	102
A.3. Methodenbeschreibungen	104
A.3.1. Contextual Inquiry	105
A.3.2. Experience Sampling Method	105
A.3.3. UX Curve	105
A.3.4. Private Camera Conversation	106
A.3.5. User Data Logging	106
A.4. Beispiele für Methodensteckbriefe	106
A.4.1. Methodensteckbrief: Heuristische Bewertung	106
A.4.2. Methodensteckbrief: Interview	108
Erklärung zur Urheberschaft	111
Erklärung zur Lizenzierung und Publikation dieser Arbeit	112
Inhalt des beigelegten Datenträgers	114

Abbildungsverzeichnis

1.	Parteivortrag im Zivilprozess: Vom Schriftsatztausch zum digitalen Basisdokument	8
2.	Erster Designvorschlag für das digitale Basisdokument	15
3.	Funktionaler Prototyp des digitalen Basisdokuments	16
4.	Basisdokumentprototyp im Reallabor Strukturvorgaben	17
5.	Abhängigkeit von Gestaltungsaktivitäten im UCD-Prozess	49
6.	Ergebnisse der Analyse der Eignung von UX-Methoden für das Reallabor Strukturvorgaben	71

Zusammenfassung

Im Rahmen eines transdisziplinären Digitalisierungsprojekts der Justiz („Strukturvorgaben für den Parteivortrag“) wird aktuell eine prototypische Anwendung des sogenannten Basisdokuments erprobt. Die Erprobung findet in Form eines Reallabors statt. In dieser Arbeit wurde untersucht, wie Evaluationsmethoden der Mensch-Maschine-Interaktion auf die Situation im Reallabor angewandt werden können. Dafür wurde im ersten Schritt das Forschungsprojekt formal (zwischen-)evaluiert, die Wahrnehmung des Basisdokuments in der juristischen Fachwelt analysiert und die Methode Reallabor untersucht und eingeordnet. Im zweiten Schritt wurde der Forschungsstand von Evaluationsmethoden der Mensch-Maschine-Interaktion aufgearbeitet und deren Anwendung auf das Reallabor mit seinen spezifischen Rahmenbedingungen analysiert. Daraus wird einerseits eine konkrete Empfehlung für die Evaluation des Basisdokument-Reallabors erarbeitet und diskutiert, andererseits soll die erarbeitete Analyse der Evaluationsmethoden auch für andere Reallabore eine Orientierungshilfe darstellen. Dazu dient der erstellte Kriterienkatalog, die Sammlung von Evaluationsmethoden und eine Übersicht über die Eignung der Methoden zur Anwendung im Reallabor.

Abstract

A prototype application of the so-called "Basisdokument" is currently being tested as part of a transdisciplinary digitization project of the German judiciary. Testing is taking place in the form of a regulatory sandbox. This work investigated how evaluation methods of human-computer interaction can be applied to the situation in the regulatory sandbox. In the first step, we conducted a (preliminary) evaluation of the research project, analyzed the perception of the "Basisdokument" in the legal community and examined the regulatory sandbox method. In the second step, we reviewed the state of research on evaluation methods of human-computer interaction and analyzed their application to the regulatory sandbox situation with its specific conditions. On the one hand, a concrete recommendation for the evalu-

ation of the "Basisdokument" regulatory sandbox will be developed and discussed; on the other hand, the analysis of the evaluation methods developed should also provide guidance for other real-world laboratories. The catalog of criteria, the collection of evaluation methods and an overview of the suitability of the methods for use in the regulatory sandbox serve this purpose.

Aufgabenstellung

Im Rahmen des interdisziplinären Reallabors „Strukturvorgaben für den Parteivortrag im Zivilprozesses“ wird aktuell das sogenannte Basisdokument im Gerichtsbetrieb erprobt. Als Ergebnis des Projekts soll eine empirisch fundierte Empfehlung über die Einführung des Basisdokuments durch den Gesetzgeber entstehen. Das Erkenntnisinteresse ist vorrangig auf die zukünftige Rechtssetzung gerichtet. Ergänzend zum regulatorischen Lernprozess sollen aber auch Erkenntnisse für die technische Umsetzung gewonnen werden. Ziel dieser Arbeit ist die Analyse der Anwendung von Evaluationsmethoden der Mensch-Maschine Interaktion für das Projekt im Rahmen des Reallabors. Dabei soll zum einen das Forschungsprojekt als solches (zwischen-)evaluiert werden, als auch begründeter Vorschlag für eine Evaluationsmethodik erarbeitet werden. Im Fokus steht dabei die besondere Reallaborsituation, die eine hohe ökologische Validität¹ aufweist, die Planung einer Evaluation aber schwierig macht.

Konkrete Schritte zum Erreichen dieses Ziels sind die Einarbeitung in das Forschungsprojekt, die Durchführung einer formativen Evaluation des aktuellen Projektstandes, sowie eine Analyse der Wahrnehmung des Basisdokuments und dessen Erprobung in der juristischen Fachwelt. Außerdem soll die Methode „Reallabor“ untersucht und im internationalen Vergleich solcher Ansätze eingeordnet werden. Zur Planung der Evaluation arbeite ich den Forschungsstand von Evaluationsmethodik im Fachbereich der Mensch-Maschine-Interaktion (*Human Computer Interaction*, HCI bzw. MMI) auf, um dann eine Analyse der Anwendung dieser Methoden auf die besondere Situation im Reallabor vorzunehmen. Ich erarbeite eine Empfehlung für eine Methodenkombination und diskutiere meine Ergebnisse.

¹Ökologische Validität bezieht sich auf die empirische Gültigkeit eines Studienergebnisses für das Alltagsgeschehen, gemeint ist die Übertragbarkeit und Anwendbarkeit der gewonnenen Aussagen auf andere Personen und vor allem auch andere Situationen (Generalisierbarkeit).

1. Einleitung

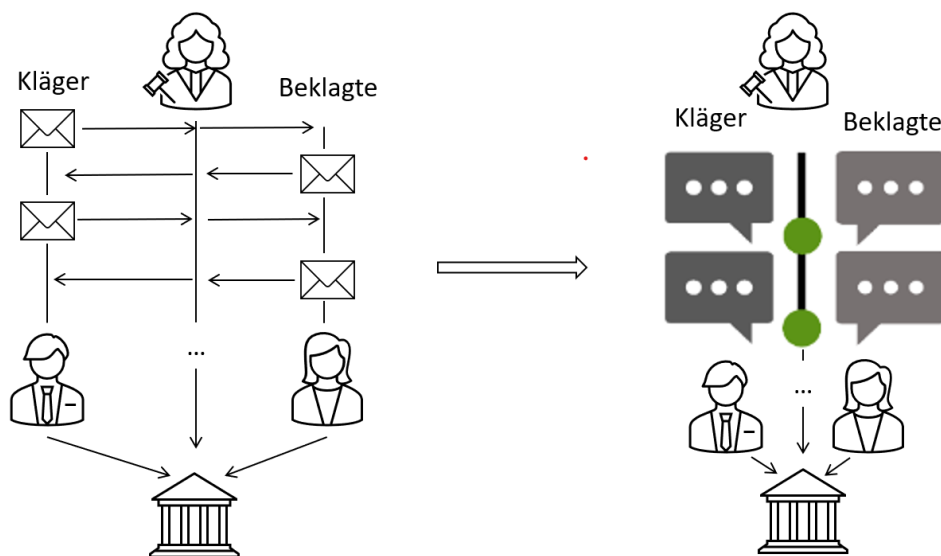


Abbildung 1: Parteilichkeit im Zivilprozess: Vom Schriftsatzverkehr (links) zum digitalen Basisdokument (rechts). Eigene Abbildung mit Logo des Projekts.

Im deutschen Zivilprozess tragen Klage- und Beklagtenpartei in Form von Schriftsätzen vor. Das Vorbringen oder der Vortrag einer Prozesspartei (Parteilichkeit) stellt die Gesamtheit der Behauptungen dar, die eine Partei in einem Zivilprozess vorbringt. Für das Zivilgericht bildet der Tatsachenvortrag der Parteien die Grundlage für das Verfahren. Es schickt also eine Seite die Klage an das Gericht, die Klage wird zugestellt, die andere Seite reagiert mit einer Klageerwidern, darauf kann es eine Replik geben und so weiter ... Abschluss dieses Schriftsatzverkehrs bildet häufig eine mündliche Verhandlung, in der der Streit verbindlich geklärt werden soll. Da es in Deutschland kaum Vorgaben zu diesem Schriftsatzverkehr (z.B. zur Länge oder Anzahl der Schriftsätze) gibt, kann dieser Vortrag sehr ineffizient und unübersichtlich werden. Richter:innen stehen dann vor der Herausforderung, den relevanten Vortrag aus den Schriftsätzen zu erarbeiten. Bei steigender Komplexität, steigender Zahl der Schriftsätze oder Wiederholungen innerhalb dieser Schriftsät-

1. Einleitung

ze besteht die Gefahr, dass relevanter Vortrag unbeachtet bleibt. Auch wenn das nicht geschieht: Die Sachverhaltsaufbereitung auf diese Art ist zeitaufwändig, ressourcenintensiv und damit gegebenenfalls auch ineffizient. Eine stärkere Strukturierung des Parteivortragens könnte diesen Prozess effizienter gestalten. Andere Länder haben dafür verschiedene Regelungen gefunden, z.B. können in Österreich Schriftsätze zurückgewiesen werden, wenn sie verworrenen sind, unklare, sinn- oder zwecklose Ausführungen oder Wiederholung enthalten; in Frankreich muss der Streitstoff nach einzelnen Behauptungen geordnet sein und in den verfahrensabschließenden Schriftsätzen muss eine Zusammenfassung des bisherigen Parteivortrags enthalten sein; in England sind die Schriftsätze in Anzahl und Länge beschränkt, Tatsachenbehauptungen sind chronologisch abschnittsweise zu gliedern (Mielke & Wolff, 2022a). In Deutschland werden verschiedene Ansätze für eine bessere Strukturierung in der juristischen Fachwelt schon seit Anfang der 1990er Jahre diskutiert (siehe z.B. Schwarz (1992); Gaier (2013); Zwickel (2016); Vorwerk (2017); Köbler (2018); Effer-Uhe (2018); Greger (2019) oder eine Übersicht bei Mielke & Wolff (2022a)).

Die Arbeitsgruppe „Modernisierung des Zivilprozesses“ (2021) schlägt als Lösungsansatz ein digitales „Basisdokument“ vor, in dem beide Parteien ihren Vortrag gesammelt und strukturiert vorbringen. Damit soll zum einen die Strukturierung des Parteivortrags verbessert bzw. unterstützt werden, zum anderen sollen die Chancen der Digitalisierung genutzt werden. Hintergrund der vorliegenden Arbeit bildet ein transdisziplinäres Projekt zur Digitalisierung im Zivilprozess, das auf diesem Vorschlag basiert. Im Forschungsprojekt „Strukturvorgaben für den Parteivortrag“ der Justizministerien Bayern und Niedersachsen wird aktuell ein Prototyp zur Strukturierung in Form eines Reallabors erprobt. Statt des Vortrags in Schriftsatzform wird also in realen Verfahren an vier Gerichten der Basisdokumentprototyp eingesetzt. Somit sind an der Erprobung Richter:innen und Anwält:innen sowie deren Angestellte beteiligt. Bis Dezember 2023 wurden von Richter:innen etwa 50 potenziell geeignete Verfahren gemeldet. Tatsächlich eingesetzt wird der Basisdokumentprototyp aber nur bei wenigen Verfahren (niedriger einstelliger Bereich),

weil eine oder beide Parteien die Nutzung ablehnen. Neben einem regulatorischen Erkenntnisinteresse (Was können Gesetzgeber:innen für die zukünftige Rechtssetzung lernen?) sollen im Projekt auch Erkenntnisse für die technische Umsetzung gewonnen werden. Um die angestrebte empirisch fundierte Empfehlung über die Einführung des Basisdokuments durch Gesetzgeber:innen zu erreichen, muss die Erprobung evaluiert werden. Eine solche Evaluation ist unter anderem wichtig, damit die Ergebnisse tatsächlich in der Politik wahrgenommen und Empfehlungen umgesetzt werden (Teirlinck et al., 2013, S. 368). Die besondere Situation im Reallabor gibt einige Rahmenbedingungen für die Evaluation vor und erschwert zum Teil die Anwendung typischer Evaluationsmethoden. Der Fokus der vorliegenden Arbeit liegt auf der Frage, welche Evaluationsmethoden der HCI-Forschung sich auf das transdisziplinäre Reallabor anwenden lassen.

Einzelne Ziele meiner Arbeit sind: (a) einen vollständigen Überblick über das Forschungsprojekt „Strukturvorgaben für den Parteivortrag“ zu geben und eine Projektanalyse zu liefern, die Probleme in der Planung und Umsetzung des Reallabors aufzeigt, (b) die Forschungsmethode *Reallabor* zu erklären und in ihren internationalen Kontext und ihren Kontext als politisches Instrument einzuordnen, (c) eine Übersicht über die gängigen HCI-Forschungsmethoden zu geben, mit Fokus auf Evaluationsmethoden, die im Feld angewandt werden (können), (d) eine Analyse der Anwendung dieser Evaluationsmethoden auf die besondere Situation im Reallabor vorzunehmen und (e) eine Empfehlung für die Methodik der Evaluation im vorliegenden Projekt zu geben. So bietet (g) die erarbeitete Analyse der Evaluationsmethoden auch für andere Reallabore eine Orientierungshilfe.

Meine Arbeit soll zur Reflexion des Projekts beitragen und einen Leitfadens für die Analyse und Bewertung der Eignung von Evaluationsmethoden darstellen – auch über den Kontext des Reallabors Strukturvorgaben hinaus für andere transdisziplinäre Projekte. Die erarbeitete Sammlung von Bewertungskriterien und Evaluationsmethoden kann in anderen Projekten genutzt werden bzw. als Ausgangspunkt für eine Bewertung dienen. Auch das kondensierte Wissen zur Methode Reallabor

und den Forschungsmethoden der HCI kann für andere Projekte nützlich sein, besonders womöglich für andere Reallabore, an denen keine Forschenden der HCI beteiligt sind, die aber ebenfalls innovative Technologien erproben und evaluieren.

Im Weiteren ist die Arbeit wie folgt aufgebaut: In Kapitel 2 wird die Entstehung und der Verlauf des Forschungsprojekts „Strukturvorgaben für den Parteivortrag“ an der Universität Regensburg geschildert. Da die Erprobung in Form eines Reallabors durchgeführt wird, gibt Kapitel 3 einen Überblick über die internationale Entwicklung, sowie die Entstehung von Reallaboren in Deutschland, Vorgaben zur Evaluation in Reallaboren und die Rahmenbedingungen des vorliegenden Reallabors „Strukturvorgaben für den Parteivortrag“ (im Folgenden als *Reallabor Strukturvorgaben* bezeichnet). In Kapitel 4 ist dargestellt, wie die juristische Fachwelt auf die Idee des Basisdokuments und die Erprobung im Reallabor reagiert und wie der im Reallabor eingesetzte Prototyp wahrgenommen wird. Kapitel 5 gibt einen detaillierten Überblick über den aktuellen Projektstand, die Ziele des Projekts, aktuelle Probleme im Reallabor und Erwartungen an die Evaluation. In Kapitel 6 werden die grundlegenden Forschungsmethoden der HCI nach Lazar et al. (2017) vorgestellt, in Kapitel 7 die Rolle der Feldforschung in der HCI genauer beleuchtet und im Feld eingesetzte Evaluationsmethoden geschildert. Die Planungen der Projektgruppe zur Evaluation sind in Kapitel 8 festgehalten. Die Analyse der Anwendung bzw. des Einsatzes der HCI-Methoden im Reallabor Strukturvorgaben wird dann in Kapitel 9 beschrieben: zum einen der Kriterienkatalog, der die Rahmenbedingungen des Reallabors Strukturvorgaben und die Ziele der Evaluation enthält, zum anderen die Ergebnisse der Analyse und deren Diskussion mit einer Empfehlung der einzusetzenden Methodenkombination. In Kapitel 10 folgen Diskussion und Fazit, Kapitel 11 schließt die Arbeit mit einem Ausblick ab.

2. Projektverlauf und Forschungsauftrag „Strukturvorgaben für den Parteivortrag“

Im Folgenden werden Hintergrund, Kontext und Verlauf des Projekts „Strukturvorgaben für den Parteivortrag“ beschrieben. Dabei gehe ich zuerst auf die Problematik der fehlenden Strukturierung des Parteivortrags im Zivilprozess ein und stelle das digitale Basisdokument als Lösungsvorschlag vor. Die Beteiligung des Lehrstuhls Medieninformatik seit 2021 zur Erarbeitung einer prototypischen Umsetzung des Basisdokuments wird ebenfalls dargestellt. Als neueste Entwicklung im Projekt stelle ich zuletzt die Planung und Rahmenbedingungen der Erprobung des Basisdokumentprototyps im Reallabor Strukturvorgaben vor. Abschließend findet sich die Abgrenzung einiger zentraler Begrifflichkeiten der vorliegenden Arbeit.

Grundlage für das nachfolgend beschriebene Pilotprojekt bildet das Diskussionspapier der Arbeitsgruppe „Modernisierung des Zivilprozesses“ (2021), das den Vorschlag für ein digitales Basisdokument als Strukturierungsmöglichkeit enthält (Mielke & Wolff, 2022b).

Ziel der Strukturierung ist die gemeinsame Erarbeitung des entscheidungserheblichen Sachverhalts durch die Parteien und das Gericht nach bestimmten Vorgaben, um so „den Zivilprozess insgesamt effektiver, schneller und qualitativ besser führen [zu] können“ (Arbeitsgruppe „Modernisierung des Zivilprozesses“, 2021, S. 33). Ziel der Arbeitsgruppe war eine umfassende Untersuchung dazu, „wie neue technische Möglichkeiten im Zivilprozess sinnvoll nutzbar gemacht werden können, um Gerichtsverfahren bürgerfreundlicher, effizienter und ressourcenschonender zu gestalten“ (Arbeitsgruppe „Modernisierung des Zivilprozesses“, 2021, S. 2). Aktuell wird der Sachverhalt von den Parteien in Schriftsätzen vorgebracht und muss von Richter:innen aus diesen Schriftsätzen erarbeitet bzw. herausgefiltert werden. Dieser Vortrag in Schriftsätzen ist im internationalen Vergleich in Deutschland wenig

2. Projektverlauf und Forschungsauftrag „Strukturvorgaben für den Parteivortrag“

reguliert (Mielke & Wolff, 2022a, S. 4 ff.) und führt nach Ansicht der Arbeitsgruppe zu Problemen: Gerade bei steigender Komplexität des Geschehens, steigender Zahl der Schriftsätze und Wiederholungen innerhalb der Schriftsätze besteht die Gefahr, dass relevanter Vortrag unbeachtet bliebe (also „übersehen“ wird); in jedem Fall sei die Sachverhaltsaufbereitung auf diese Art zeitaufwendig und ressourcenintensiv (Arbeitsgruppe „Modernisierung des Zivilprozesses“, 2021, S. 31 ff.).

Laienhaft gesprochen, lässt sich der Vorgang im Zivilprozess wie folgt beschreiben: Die Klagepartei (bzw. deren Anwält:innen) verfasst die Klage und stellt sie (digital oder per Post) an das Gericht zu, das Gericht übermittelt der Beklagtenpartei (bzw. deren Anwält:innen) die Klage, die Beklagtenpartei reagiert darauf mit einer Klageerwiderung, die Klagepartei kann darauf mit einer Replik reagieren und so weiter... Die Schriftsätze sind in ihrer Anzahl, Länge und Form grundsätzlich nicht beschränkt (Mielke & Wolff, 2022a, S. 3). Zur Verhandlung erarbeiten Richter:innen den relevanten Sachverhalt aus diesen Schriftsätzen.

Diese Problematik der Sachverhaltsaufbereitung wird in der juristischen Fachwelt schon länger diskutiert und vor allem in den letzten Jahren wurden verschiedene Vorschläge zur Strukturierung im Zivilprozess gemacht (siehe z.B. Schwarz (1992); Gaier (2013); Zwickel (2016); Vorwerk (2017); Köbler (2018); Effer-Uhe (2018); Greger (2019)). Eine inhaltliche Übersicht dieser Vorschläge ist bei Mielke & Wolff (2022a, S. 6 ff.) zu finden.

Um die Diskussion durch Einsichten einer praktischen Erprobung zu ergänzen, bildete sich an der Universität Regensburg Anfang 2021 in einem Pilotprojekt eine interdisziplinäre Arbeitsgruppe aus Rechtswissenschaftler:innen und Medieninformatiker:innen. Ziel dieses Pilotprojekts ist die Erprobung und Fortentwicklung des im Diskussionspapier entwickelten Basisdokuments als Instrument zur Strukturierung von Parteivortrag im Zivilprozess.² Sowohl für Richter:innen als für Anwält:innen soll das Basisdokument letztlich Arbeitserleichterung bedeuten. Für Richter:innen geht es dabei hauptsächlich um einen besseren Überblick durch die

²Einen Überblick über die Arbeit der Projektgruppe seit Projektbeginn kann man auf der Homepage gewinnen: <https://www.uni-regensburg.de/forschung/reallabor-parteeivortrag-im-zivilprozess/das-forschungsprojekt/index.html>, abgerufen am 22.12.2023

2. Projektverlauf und Forschungsauftrag „Strukturvorgaben für den Parteivortrag“

Strukturierung. Aktuell ist ihre Arbeit stark davon geprägt, „aus den Schriftsätzen der Parteien die relevanten Tatsachen herauszufiltern und das Streitige vom Unstreitigen zu trennen“ (Arbeitsgruppe „Modernisierung des Zivilprozesses“, 2021, S. 31). In der Praxis bedeutet das für viele Richter:innen, dass sie den Kläger- und Beklagtenvortrag in einer Relationstabelle nebeneinander darstellen und dadurch einen Überblick über den Prozessstoff erarbeiten. Das Basisdokument soll die Erstellung einer Relationstabelle überflüssig machen, indem direkt im Dokument die Beiträge in Bezug zueinander gesetzt werden. Auch für Anwäl:innen soll das Basisdokument Vorteile gegenüber herkömmlichen Schriftsätzen bringen: Der Prozessstoff ist übersichtlich dargestellt und es ist klar ersichtlich, auf welche Aussagen der Gegenseite schon reagiert wurde. Neue Beiträge enthalten dann (im Gegensatz zu neuen Schriftsätzen) immer neuen Vortrag, sodass insgesamt weniger Wiederholungen auftauchen und der Parteivortrag potenziell verschlankt ist. Im Optimalfall führt das auch zu insgesamt kürzeren Verfahrenszeiten. Ganz praktisch kann das gemeinsame Basisdokument auch in der mündlichen Verhandlung als Entscheidungsgrundlage genutzt werden.

Im Sommer 2021 wurden im Praxisseminar im Masterstudiengang Medieninformatik von Studierenden mithilfe nutzerzentrierter Entwicklungsmethodik (*User Centered Design*, siehe Kapitel 6.1) und Schwerpunkt auf der zielgruppenspezifischen Erhebung der Anforderungen von Anwäl:innen und Richter:innen erste prototypische Vorschläge erstellt, die die mögliche Arbeitsweise innerhalb einer Basisdokument-Anwendung visualisieren (Mielke & Wolff, 2022b; Hahn et al., 2021; Freisleben et al., 2021). Ein erster Prototyp, an dessen Erstellung ich als Teilnehmerin des Praxisseminars beteiligt war, ist in Abbildung 2 zu sehen. Die entstandenen Prototypen wurden auf einem Staatsempfang des bayerischen Justizministers im Oktober 2021 anhand eines Projektvideos³ vorgestellt.

Anhand dieses Videos wurde eine qualitative Studie zur Technikakzeptanz der Basisdokument-Anwendung bei Rechtsanwäl:innen durchgeführt (Böhm et al., 2023).

³Link zum Video auf der Projekthomepage: <https://www.uni-regensburg.de/forschung/reallabor-parteeivortrag-im-zivilprozess/das-vorprojekt/index.html>, abgerufen am 20.06.2023

2. Projektverlauf und Forschungsauftrag „Strukturvorgaben für den Parteivortrag“

The screenshot displays a digital legal document interface. At the top, it shows the case number '10137/15 - Parteivortrag' and user information 'ich, katrinrichter'. The main content is organized into several sections:

- Klagepartei (Plaintiff):** Firma Industri, Industriberatung GmbH, represented by Karl Kunz.
- Beklagtenpartei (Defendant):** Oswald Brack, represented by Fritz Bltner.
- Streitgrund und Schadensersatz:** Details on the legal basis and damages claimed.
- Verteidigungsanträge:** Defence applications filed by the defendant.
- Anspruchsbegründender Sachverhalt:** The facts justifying the claim, including a detailed description of a traffic accident involving a car and a motorcyclist.

The interface includes a search bar, navigation tabs, and a right-hand sidebar with 'Anmerkungen' and 'Entscheidungsgründe'.

Abbildung 2: Erster Designvorschlag für das digitale Basisdokument aus der richterlichen Perspektive. Prototyp von Hahn et al. (2021). Abbildung aus dem Projektbericht.

Als kritischster Faktor seien hierbei die „(wahrgenommene) Einschränkung der anwaltlichen Freiheit bzw. des Einsatzes taktischer Mittel“ zu bewerten (S. 8). Der Prototyp wurde jedoch als „bedienbar“ und „nützlich“ wahrgenommen, möglichen Mehrwert könnten kürzere Schriftsätze, gesteigerte Effizienz und Redundanzvermeidung bieten (S. 8). Aus den Ergebnissen dieser Studie wurden 26 Kategorien ermittelt, die in die Evaluation des Reallabors einfließen sollen, sodass die erhobenen qualitativen Daten womöglich durch quantitative Erhebungen ergänzt werden können (S. 9). Diese Kategorien beziehen sich auf den Nutzungskontext (11 Nennungen), Feedback zum Prototypen (5 Nennungen), Akzeptanz des Basisdokuments (6 Nennungen) und Konzept (4 Nennungen) (S. 5).

Eine weitere Gruppe Medieninformatik-Masterstudierender setzte im Praxisseminar im Sommer 2022 in einem iterativen Prozess einen funktionsfähigen Prototyp der Basisdokument-Anwendung als Web-Anwendung um (Bink et al., 2022). Ein Screenshot der Anwendung ist in Abbildung 3 zu sehen.

Dieser Prototyp wird seit 2023 von Mitarbeitenden am Lehrstuhl Medieninformatik kontinuierlich weiterentwickelt und ist unter <https://app.parteivortrag.de/> zu erreichen. Ein aktueller Screenshot der Anwendung mit beispielhaften Beiträgen ist in Abbildung 4 zu sehen. Die Anwendung dient im Rahmen des Realla-

2. Projektverlauf und Forschungsauftrag „Strukturvorgaben für den Parteivortrag“

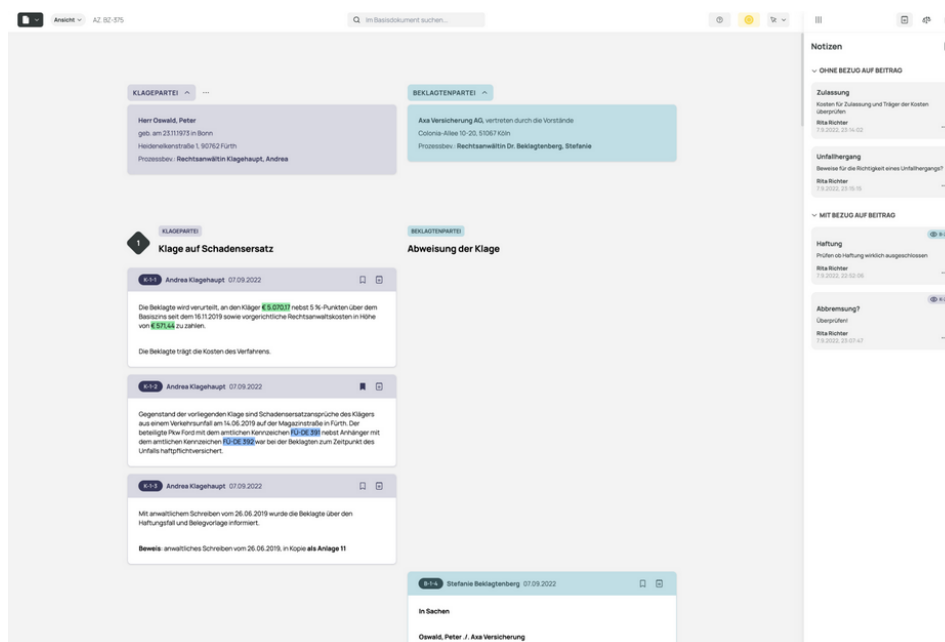


Abbildung 3: Funktionaler Prototyp des digitalen Basisdokuments von Bink et al. (2022). Screenshot aus dem Projekt-Wiki.

bors „Strukturvorgaben für den Parteivortrag im Zivilprozess“ der praktischen Erprobung des Konzepts eines digitalen Basisdokuments.

In Form eines Reallabors soll der Basisdokumentprototyp unter realen Bedingungen in der gerichtlichen Praxis an vier Gerichten in Bayern und Niedersachsen von Anwält:innen und Richter:innen erprobt werden. Die Rahmenbedingungen dafür sind in einem Forschungsvertrag zwischen den Justizministerien der beiden Länder und der Universität Regensburg festgehalten. Forschungsziel ist die wissenschaftliche Evaluation der Frage, „ob Softwarelösungen denkbar sind, mit denen der herkömmliche Austausch von Schriftsätzen im Zivilprozess durch ein strukturiertes Format ersetzt werden kann“ (Forschungsvertrag, 2022, S. 2). Anhand des Basisdokumentprototyps soll in ausgewählten Zivilverfahren erprobt werden, ob sich die Arbeit damit als vorteilhaft erweist. Die Erprobung ist ergebnisoffen geplant, am Ende des Projekts soll „eine empirisch fundierte Empfehlung über die Einführung des Basisdokuments durch den Gesetzgeber stehen. [...] Das Erkenntnisinteresse ist vorrangig auf die zukünftige Rechtssetzung gerichtet. Ergänzend zum regulatorischen Lernprozess sollen aber auch Erkenntnisse für die technische Umsetzung gewonnen werden“ (Forschungsvertrag, 2022, S. 13). Den methodischen Rahmen

2. Projektverlauf und Forschungsauftrag „Strukturvorgaben für den Parteivortrag“

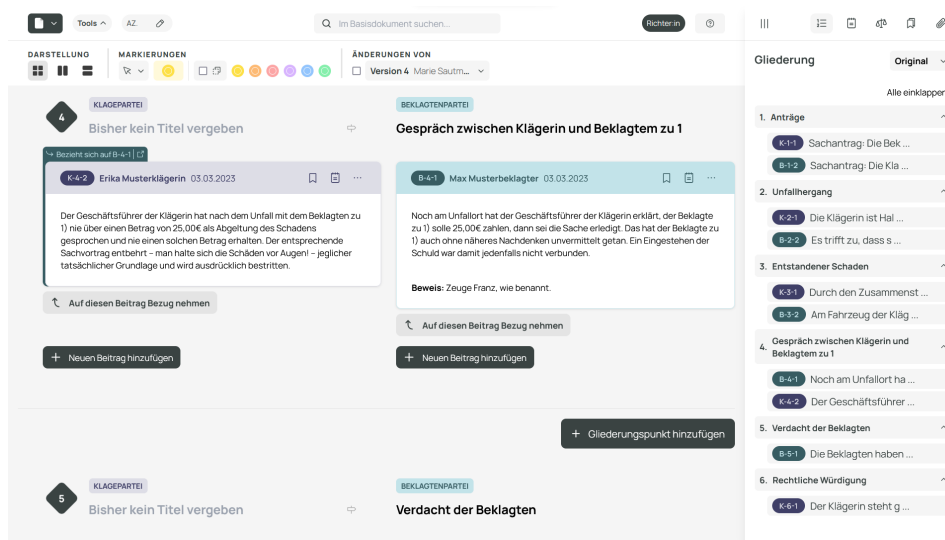


Abbildung 4: Basisdokumentprototyp im Reallabor Strukturvorgaben, Stand Dezember 2023. Screenshot der Anwendung mit beispielhaften Inhalten.

soll der Ansatz des *User Centered Design* bilden (Forschungsvertrag, 2022, S. 2).

Der aufgestellte Zeitplan sah ursprünglich eine Vorbereitungsphase bis Ende 2022, danach eine Erprobungsphase bis Ende 2023 (im Juli 2023 wurde eine Verlängerung der Erprobungsphase um sechs Monate bis Juni 2024 beschlossen) und schließlich eine Auswertungsphase bis Mitte 2024 vor. Die Vorbereitungsphase umfasst die iterative Weiterentwicklung des Basisdokumentprototyps (in Kompatibilität mit dem gegebenen Rechtsrahmen), Klärung des Einsatzbereichs und der Umsetzung im Verfahren, sowie die Gewinnung von erprobungswilligen Richter:innen und Anwält:innen. Die Erprobungsphase sieht die Erprobung des Basisdokumentprototyps an Landgerichten in erster Instanz vor und erlaubt die weitere Anpassung des Prototyps, um Feedback von Nutzer:innen integrieren zu können. In der Auswertungsphase erfolgt die Auswertung der Ergebnisse der Erprobungsphase und die Erstellung eines Abschlussberichts. Dieser Abschlussbericht „soll politische Handlungsempfehlungen enthalten und als Anstoß für eine breite Diskussion der Ergebnisse dienen“ (Forschungsvertrag, 2022, S. 16).

Die Rahmenbedingungen der Erprobung in Form eines Reallabors werden in Kapitel 3.4 detailliert beschrieben.

2. Projektverlauf und Forschungsauftrag „Strukturvorgaben für den Parteivortrag“

Zum besseren Verständnis möchte ich an dieser Stelle noch einige zentrale Begrifflichkeiten der vorliegenden Arbeit voneinander abgrenzen. Grundlage des Projekts bildet das Diskussionspapier der Arbeitsgruppe „Modernisierung des Zivilprozesses“ (2021), das den Vorschlag für ein digitales Basisdokument als Strukturierungsmöglichkeit enthält. Auf diesem Diskussionspapier baut die Arbeit im Projekt „Strukturvorgaben für den Parteivortrag“ auf, das im Forschungsauftrag der beiden beteiligten Justizministerien festgeschrieben ist. Innerhalb dieses Projekts „Strukturvorgaben für den Parteivortrag“ möchte ich die Idee bzw. das Konzept des Basisdokuments von dem Basisdokumentprototyp unterscheiden. Das Konzept meint das Basisdokument, wie es im Diskussionspapier der Arbeitsgruppe „Modernisierung des Zivilprozesses“ (2021) vorgeschlagen ist – der Prototyp ist eine konkrete Umsetzung dieses Vorschlags, er bildet dabei nur eine Möglichkeit der Umsetzung ab (auch andere sind denkbar). Dieser konkrete Basisdokumentprototyp ist es auch, der im Reallabor „Strukturvorgaben für den Parteivortrag“ (im Folgenden auch verkürzt als „Reallabor Strukturvorgaben“ bezeichnet) praktisch erprobt wird. Das Reallabor Strukturvorgaben stellt dabei nur einen Teil des Projekts „Strukturvorgaben für den Parteivortrag“ dar, die Beschreibung des Projektverlaufs macht das deutlich: Die Vorarbeiten zum Reallabor sind für das Projekt essenziell. Betrachtet man die Erprobung des Basisdokumentprototyps im Reallabor Strukturvorgaben, muss eine weitere Unterscheidung vorgenommen werden: die Arbeit mit den Basisdokument-Dateien und die Arbeit mit dem Basisdokumentprototyp. Die Basisdokument-Dateien werden über den elektronischen Rechtsverkehr versandt, sodass auch Personen mit diesen Dateien arbeiten, die keine Berührungspunkte mit dem Basisdokumentprototyp haben (z.B. Angestellte in den Geschäftsstellen der Zivilkammern). Im Rahmen der Erprobung wird der Basisdokumentprototyp evaluiert – einige Einsichten können womöglich auch zum Konzept des Basisdokuments gewonnen werden, die Rückmeldungen der Nutzer:innen werden sich aber auf die Arbeit mit dem Prototyp stützen.

Nach dieser Zusammenfassung des Projektverlaufs und begrifflichen Einordnung wird im nächsten Kapitel die Forschungsmethode Reallabor beschrieben.

3. Die Methode Reallabor

Im Folgenden soll die Methode Reallabor vorgestellt werden: ihre internationale Entwicklung, ihre Entstehung und Ausprägung als politisches Instrument in Deutschland und die deutsche Reallaborlandschaft. In einem Exkurs wird die Digitalisierung in der deutschen und bayrischen Justiz kommentiert. Schließlich werden die spezifischen Rahmenbedingungen des Reallabors Strukturvorgaben genauer geschildert und kommentiert, was die Erprobung in Form eines Reallabors für das Projekt bedeutet.

Der Begriff „Reallabor“ sowie verwandte Konzepte wie *Living Labs*, Innovationslabore, *Sandboxes*, Experimentierräume, Realexperimente, Pilot- und Demonstrationsprojekte sind kaum einheitlich zu definieren (Projektgruppe „Reallabore“, 2018, S. 4). Gemeinsam haben diese Konzepte, dass sie Räume schaffen, in denen gezielt (und zum Teil kontrolliert) Experimente in realweltlichen Settings durchgeführt werden (Schäpke et al., 2017, S. 14). Reallabore „stellen Entwicklungsumfelder dar, in denen neue Technologien, Produkte oder Services unter realen Einsatzbedingungen gleichzeitig getestet und entwickelt werden können“ (Anduschus et al., 2023, S. 11). Zentraler Forschungsmodus von Reallaboren ist die transdisziplinäre Kooperation: in ihnen arbeiten Wissenschaftler:innen verschiedener Disziplinen mit Praxisakteur:innen zusammen, „gestalten Forschungsprojekt und Ergebnisse gemeinsam durch [Zusammenarbeit] sowie die Differenzierung und Integration unterschiedlicher Wissensbestände, Methoden und Konzepte“ (Schäpke et al., 2017, S. 5).

Die Einrichtung von Reallaboren als Experimentierräumen wurde in Deutschland 2018 im Koalitionsvertrag festgeschrieben⁴. Reallabore sind also nicht nur als Forschungsansatz zu verstehen, sondern auch als politisches Instrument. Kernele-

⁴Koalitionsvertrag 2018 – Z. 1481 ff.: „Wir wollen Open-Innovation-Ansätze, soziale Innovationen sowie inter- und transdisziplinäre Ansätze fördern und Experimentierräume einrichten, um innovative technische Systeme und neue Geschäftsmodelle zu erproben.“ https://www.bpb.de/system/files/dokument_pdf/Koalitionsvertrag_2018.pdf, abgerufen am 03.05.2023

mente von Reallaboren sind der begrenzte Erprobungs(zeit)raum, die Nutzung rechtlicher Gestaltungsspielräume und aktives regulatorisches Lernen.

Im Folgenden soll ein kurzer Überblick über die internationale Einrichtung von Reallaboren gegeben werden, gefolgt von einer Schilderung der Entstehung und Definition der Methode „Reallabor“ in Deutschland. Anhand einiger Beispiele möchte ich einen Einblick in die Reallaborlandschaft in Deutschland geben und in einem kurzen Exkurs auch die Digitalisierung in der bayrischen Justiz kommentieren. Schließlich werden ergänzend zu bisherigen Beschreibungen die Rahmenbedingungen des Reallabors „Strukturvorgaben für den Parteivortrag“ detailliert dargestellt und zuletzt die ersten Überlegungen der Projektgruppe zur Evaluation der Erprobung geschildert.

3.1. Reallabore international

Bei der Schaffung von Reallaboren als Testräume für Innovation und Regulation ist Deutschland kein Einzelfall.

Eines der ersten Reallabore entstand unter dem Namen „PlaceLab“ 1999 am Massachusetts Institute of Technology (MIT): eine vollständig mit Sensorik ausgestattete Wohnung, um das Alltagsverhalten der Menschen daheim zu untersuchen (Anduschus et al., 2023, S. 12). Nokia hat 2001 das erste europäische Reallabor ins Leben gerufen („NokiaSpaceLab“), gefolgt von weiteren städtischen Projekten, bei denen Haushalte mit neuesten Technologien ausgestattet, neue Verkehrskonzepte getestet oder neue Kommunikationstechnologien erforscht wurden (Anduschus et al., 2023, S. 13). Das 2006 gegründete *The European Network of Living Labs*⁵ (ENoLL) vernetzt in Europa und weltweit Reallabore. Dies soll den Wissensaustausch, gemeinsame Aktionen und Projektpartnerschaften erleichtern.

In den USA und Europa haben sich Reallabore seither weiter verbreitet und auch in Asien gibt es einige Projekte. In den USA werden Reallabore vor allem als Innovationsmethode verstanden, um reale Daten über Nutzer:innen, Maschinen und Umgebung zu sammeln und aufzubereiten. Sie sind häufig nur ein Teil der Inno-

⁵European Network of Living Labs: <https://enoll.org/>, abgerufen am 09.10.2023

3. Die Methode Reallabor

vationsinfrastruktur (Anduschus et al., 2023, S. 14f.), z.B. im Rahmen von *Innovation Labs* wie der Google-Ausgründung X Company⁶ oder den *Harvard Innovation Labs*⁷. Auch in Europa gibt es solche *Innovation Labs* (z.B. von SAP in Kooperation mit der Universität Duisburg-Essen⁸, Siemens⁹ oder Porsche¹⁰). Es beschäftigt sich jedoch auch der Staat intensiv mit dem Format Reallabor (Anduschus et al., 2023, S. 17). In Asien finden sich vereinzelte staatliche Initiativen (z.B. der *Seoul Innovation Park*¹¹) und einige Kooperationsprojekte (z.B. das *Trans-Urban-EU-China*-Programm¹² oder *ProGiReg*¹³), die Reallabore umsetzen und als Forschungsmethode nutzen. Das Konzept hat sich in Europa etwas weiterentwickelt und der Fokus wird verstärkt auf Partizipation und Inklusion gelegt (Anduschus et al., 2023, S. 17), auch der regulatorische Aspekt mit seinen Experimentierklauseln steht stärker im Vordergrund.

Eine Recherche von Wittig & Schimanek (2020) zu internationalen regulatorischen Ansätzen ergab 52 Reallabor-Ansätze in 25 Ländern, die auch mit Experimentierklauseln oder ähnlichen Instrumenten arbeiten. Diese Ansätze lassen sich in drei Kategorien unterteilen: die verfassungsrechtliche Ermächtigung zur Schaffung experimenteller Gesetze (in Frankreich), Gesetze für Reallabore oder Experimente unterhalb des Verfassungsrangs (z.B. in Japan oder Nordirland) und administrative Maßnahmen, die die Durchführung von Experimenten im Rahmen besonderer Genehmigungen oder der Befreiung von Genehmigungspflichten ermöglicht (z.B. für den Bereich des autonomen Fahrens in den USA, Dänemark, Schweden, Niederlande, Polen, Russland, Kanada, Italien, Schweiz, Singapur, Norwegen) (Wittig & Schimanek, 2020, S. 4f.).

⁶X Company: <https://x.company/>, abgerufen am 09.10.2023

⁷Harvard Innovation Labs: <https://innovationlabs.harvard.edu/>, abgerufen am 09.10.2023

⁸SAP Innovation Lab: <https://www.iis.wiwi.uni-due.de/praxis/sap-innovation-lab/>, abgerufen am 09.10.2023

⁹Siemens Innovation Ecosystem: <https://ecosystem.siemens.com/>, abgerufen am 09.10.2023

¹⁰Porsche Innovation Lab: <https://www.porsche-consulting.com/de/de/ueber-uns/innovation-lab>, abgerufen am 09.10.2023

¹¹Seoul Innovation Park: <https://en.innovationpark.kr/>, abgerufen am 09.10.2023

¹²TRANS-URBAN-EU-CHINA: <http://transurbaneuchina.eu>, abgerufen am 09.10.2023

¹³ProGiReg: <https://progireg.eu/>, abgerufen am 09.10.2023

3.2. Reallabore in Deutschland

Da ich im Rahmen dieser Arbeit ein konkretes Reallabor in Deutschland betrachte, möchte ich im Folgenden genauer auf die Entstehung der Reallabore in Deutschland und ihre (politisch vorgegebene) Definition eingehen. Anhand einiger Beispiele möchte ich außerdem einen Einblick in die Reallaborlandschaft in Deutschland geben.

3.2.1. Entstehung und Definition

In einem Bericht zum Thema „Wissenschaft für Nachhaltigkeit“ empfahl eine Expert:innengruppe den Aufbau von Reallaboren in Baden-Württemberg als Orte der transdisziplinären Forschung (Expertengruppe „Wissenschaft für Nachhaltigkeit“, 2013, S. 31). Damals gab es in Deutschland bereits seit Jahren verschiedene Projekte, die einige Kriterien eines Reallabors erfüllen – jedoch ohne institutionellen Rahmen, z.B. das Freiburger Pionierprojekt Vauban¹⁴ für nachhaltige Stadtentwicklung, die *InnovationCity* Bottrop¹⁵ als Modellstadt für klimagerechten Stadtumbau oder das Bioenergiedorf Jühnde¹⁶.

Im Koalitionsvertrag von 2018 wurde die Einrichtung von Reallaboren dann festgeschrieben, sodass der bislang fehlende institutionelle Rahmen geschaffen werden konnte. Dieser Entscheidung folgte eine Bedarfsanalyse, auf deren Basis die Projektgruppe „Reallabore“ (2018) ein Strategiepapier präsentierte. Die Projektgruppe liefert hier auch eine erste Definition:

„Reallabore als zeitlich und räumlich begrenzte Testräume für vorwiegend digitale Innovation und Regulierung sind ein Instrument, um konkrete Erfahrungen im Zusammenspiel von Innovation und Regulierung zu sammeln und so zu einer ‚smarten‘ Regulierung zu gelangen.“ (Projektgruppe „Reallabore“, 2018, S. 1)

¹⁴Quartiersarbeit Vauban: <https://quartiersarbeit-vauban.de/>, abgerufen am 10.10.2023

¹⁵InnovationCity Bottrop: <https://www.innovationcity-bottrop.de>, abgerufen am 10.10.2023

¹⁶Bioenergiedorf Jühnde: <http://www.bioenergiedorf.info>, abgerufen am 10.10.2023

Im „Handbuch für Reallabore“ legte das BMWi¹⁷ neben Empfehlungen und Beispielen aus der Praxis 2019 auch eine ausführlichere Definition vor:

„Reallabore (englisch: ‚regulatory sandboxes‘) bieten die besondere Chance, nicht nur über Innovationen zu lernen, sondern auch über deren rechtlichen Rahmen. Als ‚Testräume für Innovation und Regulierung‘ sind Reallabore durch drei Elemente gekennzeichnet.

1. Reallabore sind zeitlich und räumlich begrenzte Testräume, in denen innovative Technologien oder Geschäftsmodelle unter realen Bedingungen erprobt werden. Häufig sind innovative Technologien oder Geschäftsmodelle nur bedingt mit dem bestehenden Rechts- und Regelungsrahmen vereinbar. Schließlich waren sie bei der Schaffung der entsprechenden Gesetze häufig schlichtweg noch nicht absehbar. Daher:
2. Reallabore nutzen rechtliche Spielräume. Experimentierklauseln oder andere Flexibilisierungsinstrumente machen es möglich, Reallabore auch dann durchzuführen, wenn die zu erprobenden Technologien oder Geschäftsmodelle im allgemein gültigen rechtlichen Rahmen noch nicht vorgesehen sind.
3. Reallabore sind mit einem ‚regulatorischen Erkenntnisinteresse‘ verbunden. Das heißt: Nicht nur die Innovation steht im Fokus, sondern auch die Frage, was der Gesetzgeber für die zukünftige Rechtssetzung lernen kann. Nur wenn mit Reallaboren ein regulatorischer Lernprozess verknüpft ist, können sie auch zu besseren Gesetzen führen.“ (BMW, 2019, S. 7)

Bezüglich der Ergebnisse weist das Handbuch auch darauf hin, Ziele für das Projekt zu formulieren und messbar machen (BMW, 2019, S. 21). Für die Evaluation des Reallabors gibt das Handbuch folgende Empfehlungen:

„Die Evaluation, also die systematische Erfassung und Bewertung der

¹⁷Bundesministerium für Wirtschaft und Energie, seit Dezember 2021 Bundesministerium für Wirtschaft und Klimaschutz (BMWK)

relevanten Daten, Informationen, Ergebnisse und Wirkungen des Reallabors, ist wesentliche Grundlage für die Aufsicht und Steuerung. Sie soll angemessen, transparent und objektiv darüber informieren, in welchem Maße die Ziele des Reallabors erreicht wurden, und sie soll das Erkenntnisinteresse der beteiligten Partner erfüllen. In einigen Fällen fordert auch die entsprechende Experimentierklausel eine wissenschaftliche Evaluation.“ (BMW, 2019, S. 53)

Auch für die Nutzung der Ergebnisse soll ein Konzept erstellt und sichergestellt werden, dass der Gesetzgeber vom Reallabor lernen kann (BMW, 2019, S. 57).

Für die konkrete rechtliche Ausgestaltung der Reallabore als Innovationsräume legte das BMW ein Konzept für ein Reallabore-Gesetz vor (BMW, 2021). Dieses Konzept schlägt vor, dass ein Gesetz (a) „übergreifende Standards für Reallabore und Experimentierklauseln gesetzlich veranker[t]“ und (b) „neue Reallabore in wichtigen digitalen Innovationsbereichen ermöglich[t]“ (BMW, 2021, S. 2). Auch für die Erstellung von Experimentierklauseln wird im Konzept ein Schema vorgelegt. Wesentlicher Bestandteil der Klauseln sollen dann auch Regelungen zur Evaluation der Erprobung und zum Transfer der Technologie in den Regelbetrieb sein (BMW, 2021, S. 5).

3.2.2. Reallaborlandschaft in Deutschland: Ein Überblick

In Baden-Württemberg werden Reallabore seit 2015 vom Wissenschaftsministerium unterstützt. Bis 2020 wurden 14 Reallabore mit dem Fokus „Wissenschaft für Nachhaltigkeit“ und eine umfangreiche Begleitforschung in zwei Förderlinien („Reallabore - BaWü-Labs“ und „Reallabor Stadt“) geschaffen. Gemeinsam mit Akteur:innen aus Gesellschaft, Politik und Wirtschaft arbeiten Wissenschaftler:innen in einem ergebnisoffenen Prozess an zukunftsfähigen und nachhaltigen Lösungen.¹⁸

Auch deutschlandweit sind Reallabore zuerst und verstärkt im Bereich Nachhaltigkeit und Energieforschung beim Bundesministerium für Wirtschaft und Klima-

¹⁸Reallaborförderung BaWü: <https://mwk.baden-wuerttemberg.de/de/forschung/forschungspolitik/wissenschaft-fuer-nachhaltigkeit/reallabore>, abgerufen am 10.10.2023

3. Die Methode Reallabor

schutz (BMWK) geschaffen worden (Förderkonzept „Reallabore der Energiewende“¹⁹). Dort werden aktuell zwölf Reallabore in den Bereichen „Energieoptimierte Quartiere“ und „Sektorkopplung und Wasserstofftechnologien“ gefördert, der zweite Bereich wird außerdem seit 2021 vom Transferforschungsprojekt *Trans4Real*²⁰ begleitet. Deutschlandweit sind Reallabore und (interessierte) Projektpartner:innen im Netzwerk „Reallabore der Nachhaltigkeit“²¹ vernetzt.

Solche Netzwerke helfen, einen Eindruck der Reallaborlandschaft in Deutschland zu bekommen. Eine komplette Übersicht zu erarbeiten, hat sich für mich aber als schwierig herausgestellt: Es gibt sehr viele Projekte, die einige Kriterien eines Reallabors erfüllen, ohne jedoch als Reallabor betitelt zu werden. Der institutionelle Rahmen wurde bundesweit erst 2018 geschaffen und die Verbreitung der Begrifflichkeiten sowie die Einrichtung von entsprechenden Förderlinien folgen seitdem. Eine Liste aller Reallabore in Deutschland konnte ich nicht finden. Einen themenübergreifenden Einblick erlaubt jedoch der Wettbewerb „Innovationspreis Reallabore: Testräume für Innovation und Regulierung“ des BMWK²², der 2022 zum zweiten Mal stattfand. Am Wettbewerb konnten Ideen für Reallabore, laufende und abgeschlossene Reallabore teilnehmen. Aus über 100 Einreichungen wurden zehn Sieger gewählt, die im Folgenden kurz vorgestellt werden, um einen Einblick in die Diversität der deutschen Reallaborlandschaft zu geben.

Mit dem **LastMileCityLab**²³ am Innovationszentrum Efeu-Campus wird Bruchsal zum Reallabor für zukunftsweisende, urbane Gütermobilität. Im ersten Schritt werden autonom fahrende Roboter zur Paketauslieferung erprobt, im zweiten wird das Quartier für weitere Lösungen geöffnet (z.B. „mobile postoffices“, autonome Schwerlastdrohnen und autonome LKW). Ausnahmegenehmigungen sind für autonomes Fahren und Fliegen notwendig.²⁴

¹⁹Förderkonzept Reallabore der Energiewende: <https://www.energieforschung.de/foerderkonzept-reallabore>, abgerufen am 10.10.2023

²⁰Trans4Real: <https://www.ffe.de/projekte/trans4real-transferforschung-fuer-die-reallabore-der-energiewende-zu-sektorkopplung-und-wasserstoff/>, abgerufen am 10.10.2023

²¹Reallabornetzwerk: <https://www.reallabor-netzwerk.de>, abgerufen am 10.10.2023

²²Wettbewerb Innovationspreis Reallabore: <https://www.bmwk.de/Redaktion/DE/Wettbewerb/innovationspreis-reallabore.html>, abgerufen am 11.10.2023

²³LastMileCityLab: <https://efeucampus-bruchsal.de/>, abgerufen am 11.10.2023

²⁴Innovationspreis für das LastMileCityLab: <https://www.bmwk.de/Redaktion/DE/>

TeamX²⁵ (Trusted Ecosystem of Applied Medical Data eXchange) soll die Basis für eine zukunftsweisende Gesundheits- und Pflegeversorgung legen. Ziel ist die Schaffung eines Datenökosystems. Darauf aufbauend werden Ansätze erprobt, um die Kompetenz der Bürger:innen im Umgang mit ihren Daten zu erhöhen. Außerdem werden Lösungen im Bereich der Pflege- und Frauengesundheit in realer Umgebung getestet. Mit Blick auf die besonders schutzwürdigen Gesundheitsdaten sollen Weiterentwicklungsbedarfe im regulatorischen Rahmenwerk herausgearbeitet werden.²⁶

Multikopter²⁷ sind senkrechtstartende Luftfahrzeuge mit elektrisch angetriebenen Rotoren. Im Projekt soll der Einsatz von Multikoptern im Rettungsdienst bzw. in Form von Zubringern für Notärzt:innen erprobt werden. Für luftrechtliche und rettungsdienstliche Gesetze und Verordnungen bietet das Reallabor hohes Erkenntnispotenzial.²⁸

Landnetz²⁹ schafft ein digitales Experimentierfeld mit dem Fokus auf erforderliche Kommunikations- und Cloudinfrastrukturen zur Erforschung und Prüfung von Technologien zur flächendeckenden drahtlosen Datenübertragung mithilfe von 5G im ländlichen Raum, Vernetzung von landwirtschaftlichen Betrieben und die Nutzung von Datenhubs in einer ausgewählten Modellregion. Der Netzbetrieb ist aktuell dank Versuchsfunklizenzen möglich, die aufgrund des Forschungscharakters des Projekts genutzt werden können.³⁰

Das **Digitale Testfeld auf der Bundeswasserstraße Schlei** dient dem Aufbau eines europäischen Test- und Validierungszentrums für autonome maritime Systeme. Für die Erprobung steht das weltweit erste autonome und emissionsfreie Boot „ZeroOne“ des Unternehmens *Unleash Future Boats* mit Brennstoffzellen-Technologie

Wettbewerb/Fragmente/innovationspreis-reallabore-lastmilecitylab.html, abgerufen am 12.10.2023

²⁵TeamX: <https://project-team-x.eu/>, abgerufen am 11.10.2023

²⁶Innovationspreis für TeamX: <https://www.bmwk.de/Redaktion/DE/Wettbewerb/Finalisten/innovationspreis-reallabore-teamx.html>, abgerufen am 12.10.2023

²⁷Multikopter: <https://luftrettung.adac.de/multikopter/>, abgerufen am 11.10.2023

²⁸Innovationspreis für Multikopter: <https://www.bmwk.de/Redaktion/DE/Wettbewerb/Finalisten/innovationspreis-reallabore-multikopter.html>, abgerufen am 12.10.2023

²⁹Landnetz: <https://landnetz.eu/projekt/>, abgerufen am 11.10.2023

³⁰Innovationspreis für Landnetz: <https://www.bmwk.de/Redaktion/DE/Wettbewerb/Finalisten/innovationspreis-reallabore-landnetz.html>, abgerufen am 12.10.2023

zur Verfügung. Mit Blick auf den rechtlichen Rahmen erlaubt das Reallabore Erkenntnisse hinsichtlich Betriebskonzepten, Systemgrenzen und Sicherheitsbestimmungen.³¹

Tempus³² rüstet im Münchner Stadtgebiet und dem anliegenden Umland ein definiertes Testgebiet mit kommunizierender Straßeninfrastruktur aus. Es werden verschiedenste Anwendungsfälle rund um automatisierte Fahrfunktionen und innovative Mobilitätsleistungen für den Individual- und den öffentlichen Personennahverkehr erprobt, simuliert und evaluiert.³³

USpace³⁴ ist ein fest definiertes geografisches Gebiet im Hamburger Hafen, in dem unbemannte und bemannte Luftfahrt gemeinsam stattfinden kann. Auf Basis einer neuen Durchführungsverordnung der Europäischen Kommission wurde im Reallabor weltweit erstmals ein U-Space mit vollständigem Rollen- und Prozessmodell, technischen Funktionalitäten und Testflügen von Drohnen in der Praxis erprobt.³⁵

WohnZukunft sollte Möglichkeiten aufzeigen, wie mit geringem technischem und finanziellem Aufwand bis zu 30 Prozent des Energieverbrauchs in Mietwohnungen eingespart werden kann. Erprobt wurden dazu Lösungen smarterer Messtechniken und Kommunikationseinheiten. Das Reallabor versteht sich als Best-Practice-Modell für „Smart-Meter-Gateways“, das die Wirksamkeit dieser Technologie im Zeitraum 2016 bis 2020 schon belegen konnte, noch bevor der Gesetzgeber diese als Standard festgelegt hat (ab 2026 Pflicht in Mietwohnungen).³⁶

Im **Reallabor Hamburg** wurde die Mobilität von morgen in einer Metropole erprobt und darauf aufbauend ein Leitfaden für die digitale Mobilität der Zukunft

³¹Innovationspreis für das digitale Testfeld Schlei: <https://www.bmwk.de/Redaktion/DE/Wettbewerb/Finalisten/innovationspreis-reallabore-bundesswasserstra%C3%9Fschlei.html>, abgerufen am 12.10.2023

³²Tempus: <https://tempus-muenchen.de/>, abgerufen am 11.10.2023

³³Innovationspreis für Tempus: <https://www.bmwk.de/Redaktion/DE/Wettbewerb/Finalisten/innovationspreis-reallabore-tempus.html>, abgerufen am 12.10.2023

³⁴USpace: <https://droniq.de/projekte/u-space-reallabor-hamburg/>, abgerufen am 12.10.2023

³⁵Innovationspreis für USpace: <https://www.bmwk.de/Redaktion/DE/Wettbewerb/Finalisten/innovationspreis-reallabore-ospace.html>, abgerufen am 12.10.2023

³⁶Innovationspreis für Wohnzukunft: <https://www.bmwk.de/Redaktion/DE/Wettbewerb/Finalisten/innovationspreis-reallabore-wohnzukunft.html>, abgerufen am 12.10.2023

erarbeitet. Die gesellschaftliche Debatte zu digitalen Mobilitätsservices stand dabei im Zentrum, um wichtige Erkenntnisse darüber zu liefern, welche Ansätze sich in der Praxis bewähren.^{37 38}

Im **Klimaquartier Neue Weststadt**³⁹ in Esslingen wird Deutschlands erstes nahezu klimaneutrales Wasserstoffquartier mit Wohnungen, Büro- und Gewerbeflächen sowie einer Hochschule errichtet. Das Projekt zeigt, wie modernes Wohnen und Arbeiten und nachhaltige Mobilität im städtischen Kontext einen signifikanten Beitrag zum Klimaschutz leisten kann.⁴⁰

Diese zehn Projekte unterscheiden sich stark in Laufzeit (beim LastMileCityLab z.B. von 2021 bis 2027, USpace nur für 18 Monate) und Höhe der Fördermittel. Verschiedene Einrichtungen sind Mittelgeber (neben dem BMWK z.B. das Bundesministerium für Ernährung und Landwirtschaft (BMEL), der Europäischen Fonds für regionale Entwicklung (EFRE) oder das Bundesministerium für Digitales und Verkehr (BMDV)). Es sind auch nicht alle Projekte deutlich als „Reallabor“ betitelt, ebenso werden durch die meisten Förderlinien nicht nur Reallaborprojekte gefördert. Insgesamt scheint es, dass mit dem Label „Reallabor“ ein politisches Instrument geschaffen wurde, um Innovation und Regulationsevaluation zu fördern, die Projekte aber weiterhin sehr divers sind. Anduschus et al. (2023) fassen das treffend zusammen: „Während in der Theorie eingängige Definitionen des Begriffs existieren, zeigt sich in der Praxis jedoch eine erhebliche Bandbreite an Ausgestaltungen solcher Einrichtungen“ (S. 4).

3.3. [Exkurs] Digitalisierung in der (bayrischen) Justiz

Ich möchte in diesem kurzen Exkurs die Digitaloffensive der Justiz in Bayern vorstellen, um das Reallabor „Strukturvorgaben für den Parteivortrag“ weiter in seinen

³⁷RealLabHH: <https://www.bmdv.bund.de/SharedDocs/DE/Artikel/DG/AVF-projekte/reallabhh.html>, abgerufen am 12.10.2023

³⁸Innovationspreis für das Reallabor Hamburg: <https://www.bmwk.de/Redaktion/DE/Wettbewerb/Finalisten/innovationspreis-reallabore-reallaborhamburg.html>, abgerufen am 12.10.2023

³⁹Klimaquartier neue Weststadt: <https://neue-weststadt.de/>, abgerufen am 12.10.2023

⁴⁰Innovationspreis für das Klimaquartier Neue Weststadt: <https://www.bmwk.de/Redaktion/DE/Wettbewerb/Finalisten/innovationspreis-reallabore-sonderpreis.html>, abgerufen am 12.10.2023

Kontext zu setzen.

Digitalisierung in Deutschland, vor allem auch in Behörden, ist seit einigen Jahren ein wichtiges und viel diskutiertes Thema.⁴¹ In Bayern hat Justizminister Eisenreich zum April 2023 eine neue Abteilung „Digitalisierung und Innovation“ im bayerischen Justizministerium geschaffen. Die Chancen der Digitalisierung sollen auch in der bayrischen Justiz genutzt werden. Bestandteile der Offensive sind der elektronische Rechtsverkehr (an allen bayrischen Gerichten eingeführt), Videoverhandlungen (seit 2021 bei teilnehmenden Gerichten möglich), Einführung der E-Akte (muss bis 2026 deutschlandweit erfolgen), das „Legal Tech Colab“ (Inkubator und Accelerator für Start-ups im Legal-Tech-Bereich), wählbares Berufsfeld „IT-Recht und Legal Tech“ für Referendar:innen (ab Juli 2023), das Deep-Tech-Tool „Dark Web Monitor“ und Forschungs- und Pilotprojekte zur automatisierten Anonymisierung von Gerichtsurteilen.⁴²

Inzwischen reiht sich neben anderen Projekten zur Anonymisierung⁴³ oder Massenverfahren auch das Reallabor „Strukturvorgaben für den Parteivortrag“ in die Kategorie Forschungs- und Pilotprojekte ein.⁴⁴

Auf dem Digitalgipfel der Länder im November 2023 wurde außerdem beschlossen, dass der digitale Transformationsprozess in der Justiz auch die weitere Modernisierung der Prozessordnungen umfassen soll.⁴⁵ Neben den Vorhaben der Länder können auch Projekte des Bundesjustizministeriums durch die Digitalisierungsinitiative finanziert werden.⁴⁶ Hier scheint das Projekt „Strukturvorgaben für den

⁴¹Die Umsetzungsstrategie „Digitalisierung gestalten“ der Bundesregierung z.B. fußt auf der Digitalen Agenda 2014–2017 und gibt Leitlinien zur Umsetzung von digitalpolitischen Maßnahmen. Ein Digitalrat gibt der Regierung seit 2018 Impulse zur Digitalisierung.

⁴²Pressemitteilung zur bayrischen Digitaloffensive vom 05.04.2023: <https://www.justiz.bayern.de/presse-und-medien/pressemitteilungen/archiv/2023/66.php>, abgerufen am 23.10.2023

⁴³Projekt der FAU „Automatische Anonymisierung und Pseudonymisierung von Gerichtsurteilen“: <https://cris.fau.de/projects/264601690/>, abgerufen am 26.10.2023

⁴⁴siehe LinkedIn-Post zu Eisenreichs Rede beim Kongress LEGAL REVOLUTION von Christian Wolff: [https://www.linkedin.com/feed/update/urn:li:activity:7059453426048409600?updateEntityUrn=urn%3Ali%3Afs_feedUpdate%3A\(V2%2Curn%3Ali%3Aactivity%3A7059453426048409600\)](https://www.linkedin.com/feed/update/urn:li:activity:7059453426048409600?updateEntityUrn=urn%3Ali%3Afs_feedUpdate%3A(V2%2Curn%3Ali%3Aactivity%3A7059453426048409600)), abgerufen am 26.10.2023

⁴⁵Pressemitteilung zum Digitalgipfel vom BMJ: https://www.bmj.de/SharedDocs/Pressemitteilungen/DE/2023/1110_Digitalgipfel.html, abgerufen am 29.11.2023

⁴⁶Meldung zum Digitalgipfel von beck-aktuell: <https://rsw.beck.de/aktuell/daily/meldung/detail/dritter-bund-laender-digitalgipfel-beschliesst-modernisierung-der-prozessordnungen>, abgerufen am 29.11.2023

Parteivortrag“ anschlussfähig.

3.4. Reallabor „Strukturvorgaben für den Parteivortrag“

Ergänzend zur Beschreibung des Projektverlaufs (Kapitel 2) sollen im Folgenden die Rahmenbedingungen der Erprobung des Prototyps im Reallabor „Strukturvorgaben für den Parteivortrag“ (im Folgenden auch abgekürzt zu „Reallabor Strukturvorgaben“) detaillierter beschrieben werden. Diese Informationen habe ich nicht einer einzelnen Quelle entnommen, sondern im Laufe meiner Mitarbeit im Projekt in Gesprächen mit Projektmitgliedern, bei Informationsveranstaltungen am Gericht Regensburg und einem Besuch der Geschäftsstelle der 3. Zivilkammer Regensburg gesammelt.

Im Reallabor Strukturvorgaben soll der Einsatz des Basisdokumentprototyps unter realen Bedingungen in der gerichtlichen Praxis von Anwält:innen und Richter:innen in (nach Einschätzung der Richter:innen) geeigneten Zivilverfahren erprobt werden. Die Erprobung läuft seit März 2023, nach einer Verlängerung ist aktuell eine Laufzeit von 18 Monaten vorgesehen. Auch während der Erprobung erfolgen noch Anpassungen des Prototyps.

Die Erprobung findet an den Landgerichten Hannover, Landshut, Osnabrück und Regensburg in erster Instanz statt. Der Prototyp ist nicht in die vorhandene IT-Infrastruktur integriert, es gibt keine direkte Anbindung an genutzte Anwendungen (besonderes elektronisches Anwaltspostfach (beA), sowie das elektronische Integrationsportal (eIP) in Bayern und e² in Niedersachsen).

Der Basisdokumentprototyp ist als Web-Anwendung umgesetzt und wird im Browser genutzt, es werden aber keine Daten auf Servern gespeichert – es besteht keinerlei Anbindung an ein Backend. Die Anwendung erlaubt das Erstellen und Editieren eines Basisdokuments. Die Inhalte des Basisdokuments werden als PDF-Datei und als .txt-Datei heruntergeladen, als weitere .txt-Datei wird eine persönliche Bearbeitungsdatei gespeichert, die Metadaten wie Markierungen, Sortierungen oder Lesezeichen enthält. Um an einem Basisdokument zu arbeiten, lädt man im Prototyp die aktuelle Version der Basisdokument.txt-Datei und die persönliche Bearbeitungs.txt-

Datei hoch. Von den Anwält:innen werden sowohl die PDF-Datei als auch die Basisdokument.txt-Datei an das Gericht übermittelt. Die einzelnen Funktionen des Basisdokumentprototyps sind im Wiki auf der Projekthomepage erklärt.⁴⁷

Der Basisdokumentprototyp wird im vorhandenen Rechtsrahmen erprobt. Im Projekt liegt keine Experimentierklausel vor, die Ausnahmen von geltendem Recht erlaubt. Streng genommen ist das vorliegende Reallabor damit nur durch zwei der drei Elemente eines Reallabors nach der Definition des BMWi (2019) gekennzeichnet. Die Kommunikation zwischen Gericht und Parteien erfolgt daher über den Dokumentenaustausch im elektronischen Rechtsverkehr, einschließlich der Vorgaben der Elektronischer-Rechtsverkehr-Verordnung⁴⁸ (ERVV). Das rechtliche Gehör⁴⁹ und der Grundsatz des fairen Verfahrens⁵⁰ müssen gewahrt bleiben.

Durch das Fehlen einer Experimentierklausel können sowohl Richter:innen, als auch Anwält:innen nicht zur Teilnahme an der Erprobung verpflichtet werden. Richter:innen, sowie Anwält:innen und Mandant:innen beider (bzw. aller beteiligten) Parteien müssen also der Nutzung des Basisdokumentprototyps in einem Verfahren zustimmen und können ihre Zustimmung zu jedem Zeitpunkt im Verfahren widerrufen. Die Planung der Projektgruppe sieht vor, dass erprobungswillige Richter:innen gewonnen werden, die dann nach Einreichung der Klage (in Form eines traditionellen Schriftsatzes) den Parteien die Nutzung des Basisdokumentprototyps für das weitere Verfahren vorschlagen, wenn sie das Verfahren für die Erprobung geeignet finden. In einzelnen Fällen haben aber auch Anwält:innen die Klage direkt als Basisdokument-Datei eingereicht.

Die aktive Nutzung des Basisdokumentprototyps erfolgt für die Beteiligten dann unregelmäßig: Anwält:innen reagieren jeweils innerhalb gesetzter Fristen auf gegnerischen Vortrag, Richter:innen können zu verschiedenen Zeitpunkten Hinweise erteilen oder Prozessstoff aufbereiten.

⁴⁷Wiki zum Basisdokumentprototyp: <https://www.uni-regensburg.de/forschung/reallabor-informationen/wiki/index.html>, abgerufen am 22.12.2023

⁴⁸Elektronischer-Rechtsverkehr-Verordnung: <https://www.gesetze-im-internet.de/ervv/BJNR380300017.html>, abgerufen am 26.10.2023

⁴⁹Rechtliches Gehör, Grundgesetz Art. 103 Abs. 1: https://www.gesetze-im-internet.de/gg/art_103.html, abgerufen am 26.10.2023

⁵⁰Rechtsstaatsgebot, Grundgesetz Art. 20 Abs. 3: https://www.gesetze-im-internet.de/gg/art_20.html, abgerufen am 26.10.2023

3. Die Methode Reallabor

Durch die Eingliederung des Basisdokuments in reguläre Arbeitsabläufe kommen neben Anwäl:innen und Richter:innen auch weitere Angestellte der Kanzleien und Gerichte in Berührung mit dem Basisdokumentprototyp oder den Basisdokument-Dateien. Das betrifft nicht ausschließlich die Arbeit mit dem Prototyp, sondern ist teils auch durch den Datenaustausch und die damit verbundene Arbeit in Postfach und Aktenportalen bedingt.

Das Projekt wird in juristischen Kreisen mit Interesse verfolgt und kommentiert (siehe z.B. Pressemitteilungen⁵¹, Zeitschriftenbeiträge z.B. Bert (2022a); Zwickel (2023) oder Posts auf LinkedIn⁵²). Dieser Diskurs der juristischen Fachwelt wird in Kapitel 4 dargestellt.

Das Reallabor Strukturvorgaben erfüllt streng genommen nur zwei der drei Kriterien für Reallabore in Deutschland: Es ist ein zeitlich und räumlich begrenzter Testraum, in dem eine innovative Technologie unter realen Bedingungen erprobt wird (Kriterium 1) und mit einem regulatorischen Erkenntnisinteresse verbunden (Kriterium 3). Allerdings nutzt das Reallabor Strukturvorgaben keine rechtlichen Spielräume (Kriterium 2). Die Nutzung des Basisdokuments ist im allgemein gültigen rechtlichen Rahmen noch nicht vorgesehen, für die Erprobung wird aber der Basisdokumentprototyp diesen Vorgaben angepasst – nicht der rechtliche Rahmen an die zu erprobende Technologie.

⁵¹Zum Start des Reallabors: <https://www.justiz.bayern.de/presse-und-medien/pressemitteilungen/archiv/2023/52.php>, abgerufen am 27.10.2023

⁵²Einträge mit #basisdokument auf LinkedIn: <https://www.linkedin.com/feed/hashtag/?keywords=basisdokument>, abgerufen am 29.11.2023

4. Wahrnehmung des Projekts Basisdokument

Im Folgenden ist dargestellt, wie die juristische Fachwelt auf die Idee des Basisdokuments und die Erprobung im Reallabor reagiert und wie der im Reallabor eingesetzte Prototyp wahrgenommen wird. Für die Recherche habe ich mit einer gängigen Suchmaschine (DuckDuckGo) und innerhalb des Anwaltsblatts nach Beiträgen zum „Basisdokument“ gesucht und die Ergebnisse gesichtet. Pressemitteilungen und private Blogs habe ich hier nicht berücksichtigt, sonst aber alle Publikationen über das Basisdokument bzw. die Erprobung im Reallabor, die ich gefunden habe.

Diese Einordnung dient dem Verständnis der Frage, wie die Richter:innen und Anwält:innen über den Vorschlag eines Basisdokuments denken bzw. wie der Diskurs geführt wird – in Klammern ist jeweils auch der Beruf der Autor:innen festgehalten. Es wird vor allem auch deutlich, dass der Diskurs – trotz der Bemühung der Projektgruppe einen konkreten Vorschlag für eine Implementierung des digitalen Basisdokuments zu liefern – sich häufig auf die abstrakte Idee „Basisdokument“ bezieht, mit der sehr unterschiedliche Vorstellungen verbunden sind.

4.1. Wahrnehmung des Basisdokuments in der juristischen Fachwelt

Der Fokus dieser Beiträge zum Vorschlag des Basisdokuments liegt hier auf Beiträgen im deutschsprachigen Anwaltsblatt. Ausgangspunkt der vorgestellten Beiträge bildet das Diskussionspapier der Arbeitsgruppe „Modernisierung des Zivilprozesses“ (2021).

Horn (2020) (Rechtsreferendar am Kammergericht, in der Anwaltsstation) kommentiert den Vorschlag des Basisdokuments – „wenn so umgesetzt“ – als revolutionär. Er bewertet einzelne Aspekte der Idee als positiv (z.B. die Darstellung nach Art einer Relationstabelle und die gesteigerte Übersichtlichkeit). Horn merkt jedoch auch an, dass eine technische Umsetzung Zeit für flächendeckende Akzeptanz be-

4. Wahrnehmung des Projekts Basisdokument

nötige – Gewöhnung und Kennenlernen bräuchten Zeit und Energie, vor allem da im vorliegenden Fall Beteiligte mit entgegengesetzten Interessen ein gemeinsames Arbeitsprodukt schaffen müssten. Die Vorstellung schließt mit der Vermutung, dass das Basisdokument nach einer Zeit der Etablierung und Übung dabei helfen wird, „Prozesse schneller und günstiger abzuwickeln“.

Witte (2020) (Rechtsanwältin) kommentiert den Vorschlag der Arbeitsgruppe zur Strukturierung ebenfalls als besonders spannend. Es stelle sich dabei aber die Frage, „welche Auswirkungen eine solche Pflicht auf die Arbeitsweise der Anwaltschaft“ habe und ob es Nachteile für die Anwält:innen gäbe (S. 4). Witte beleuchtet die Frage aus rechtlicher Sicht, nach der Anwält:innen ein unabhängiges Organ der Rechtspflege (§ 1 BRAO) und zur Beratung und Vertretung in allen Rechtsangelegenheiten befugt sind (§ 3 BRAO). Sie stellt die Frage in den Raum, ob gesetzliche Vorgaben zur Strukturierung Anwält:innen darin beschneiden würden oder vorteilhaft wären. Für den Beibringungsgrundsatz⁵³ erscheint es Witte bedenklich, Vorgaben zum Vortrag zu machen. Als potenzielle Vorteile auch für Anwält:innen sieht Witte die Möglichkeit „konkret nachzuvollziehen, zu welchen Tatbestandsvoraussetzungen vorgetragen wurde“ und „an welchen Stellen Nachbesserungsbedarf [...] besteht“ (S. 4).

R. Köbler (2021) (Präsident des Landgerichts) bezeichnet den Vorschlag des Basisdokuments nach Vorstellung der Arbeitsgruppe als „radikal“; völlig neu dabei sei, dass „erstmal ein ganzheitlicher Vortrag entstünde“, der dem Redundanzeffekt entgegenwirke. Der Vorwurf, das Basisdokument sei insbesondere auf die Erleichterung der richterlichen Arbeit ausgerichtet, wird auch hier erwähnt – scheint jedoch hauptsächlich eine Motivation zur weiteren Diskussion zu sein.

Diesen positiven Kommentaren zum Trotz berichtet Albrecht (2021) (wissenschaftlicher Mitarbeiter am Institut für Telekommunikations-, Informations- und Medienrecht) vom Zivilrichtertag 2021, dass der Vorschlag des Basisdokuments nach Vorstellung und Diskussion mit 51 % Zustimmung nur eine knappe Mehrheit erhielt und mit 30 % Ablehnung auch entschiedene Gegner habe. Dieses Ergebnis ist

⁵³Beibringungsgrundsatz: „da mihi factum, dabo tibi ius – die Parteien liefern den Sachverhalt, das Gericht zieht daraus die rechtlichen Schlüsse“

4. Wahrnehmung des Projekts Basisdokument

nicht repräsentativ für die Justiz in Deutschland (nur etwa 80 Richter:innen beteiligten sich an der Abstimmung), soll hier aber als Stimmungsbild erwähnt werden.

Liero (2021) (Rechtsreferendarin, im DAV) berichtet von der Online-Konferenz des Forschungsinstituts für Anwaltsrecht 2021, für die juristische Fachwelt gehe es nicht mehr um das „Ob“ der Digitalisierung, sondern das „Wie“. Vorteile des Basisdokuments bestünden für das Gericht in der Ersparung der mühevollen Strukturierung und des Abfassens eines Tatbestands im Urteil. Bedenken äußerten Teilnehmer:innen bezüglich komplex aufgeladener Sachverhalte und der Frage, ob diese sich tatsächlich durch ein gemeinsames Basisdokument in den Griff bekommen ließen, sowie dem Vorgehen im Falle unstrukturierter Vorbringens innerhalb des Basisdokuments.

Bert & Windau (2021) (Rechtsanwalt und Richter) führten im ZPO-Blog (Zivilprozessordnung) eine Umfrage zu aktuell diskutierten Reformvorschlägen zum Zivilprozess durch. Die Leser:innen des Blogs sollten bewerten, welche der Themen sie am wichtigsten finden. Das Basisdokument konnte dabei nur 15,97 % der Stimmen gewinnen. Die Autoren interpretieren dieses Ergebnis dahingehend, dass das Basisdokument den 407 Teilnehmer:innen nicht als eines der „großen, das System des Zivilprozess ändernden Vorschläge“ erscheint.

Zwickel (2023) (Akademischer Oberrat an der FAU) bezeichnet das Basisdokument ganz aktuell als „wohl streitigste[n] Punkt in den Bemühungen um eine Digitalisierung des Zivilprozesses“ (S. 91). Grund hierfür sei die Befürchtung, es könne zu einer Verlagerung rechtlicher Aufgaben auf die Parteien kommen. Eine prozessuale Revolution sei es in seiner gegenwärtigen Gestalt jedoch nicht und bleibe hinter den Möglichkeiten moderner Digitaltechnik zurück. Als positive Effekte des Basisdokuments zählt Zwickel die Erleichterung der Erarbeitung des Urteilsbestands, sowie der einvernehmlichen Streitbeilegung durch kooperatives Formulieren und Verfahrensbeschleunigung durch Konzentration auf wesentlichen Streitstoff auf. Nachbesserungsbedarf sieht Zwickel bei der tatsächlichen Gewinnung digitaler Daten (im Gegensatz zur Strukturierung durch Verwendung digitaler Tools), bei der Reichweite der Strukturierung (Ausweitung auf Rechtsvortrag und Verfah-

4. Wahrnehmung des Projekts Basisdokument

ren) und bei den Strukturierungskriterien (im Basisdokument chronologisch vorgesehen).

Auch Kilian (2023) (Direktor des Soldan Instituts und des Instituts für Anwaltsrecht der Universität Köln) stellt in seinem Beitrag die Idee des Basisdokuments nach Vorstellung der Arbeitsgruppe vor. Im Rahmen des Berufsrechtsbarometers 2021 wurden 2.070 Rechtsanwält:innen zu ihrer Meinung zum Vorschlag befragt. Auch hier zeigt sich die ablehnende Haltung der Anwaltschaft (67 % Ablehnung, 33 % Zustimmung). Die ablehnende Haltung scheint sich mit den Faktoren höheres Alter, kleinere Kanzlei, rein nationale Tätigkeit zu verstärken (jedoch nicht ungewöhnlich stark). Es sei nicht ausgeschlossen, dass „ein konkreter ausgearbeiteter modifizierter Vorschlag, der Vorbehalte der Anwaltschaft adressieren würde, positiver aufgenommen wird“. Die starke Ablehnung aller Teilgruppen der Anwaltschaft spräche allerdings eher für ein grundlegendes Störgefühl gegenüber dieses Vorschlags.

Diese Beiträge zeigen, dass der Vorschlag der Arbeitsgruppe sowohl in der Literatur, als auch auf Tagungen seit 2021 kontrovers diskutiert wird. Das Basisdokument findet also durchaus Beachtung, wird jedoch auch von Richter:innen nicht nur als positiv oder gar wichtig wahrgenommen. Die Einwände gegen das Basisdokument sind vor allem aus der Anwaltschaft sehr stark und meist die gleichen.

Beteiligte der Projektgruppe haben in einer ersten Studie die Technikakzeptanz von Anwält:innen gegenüber dem Basisdokument untersucht (Böhm et al. (2023), kurz erwähnt auch schon in Kapitel 2). Das Basisdokument wurde den Anwält:innen mithilfe des Projektvideos erläutert, das die ersten Prototypen zeigt (also die Vorgänger der im Reallabor eingesetzten Software). Das Feedback zum Basisdokument fiel dabei gemischt aus. Als positive Aspekte wurden auch hier die Arbeitserleichterung für die Richterschaft, die Vermeidung unnötig langer Schriftsätze, die Steigerung der Effizienz durch Strukturierung und das Verhindern von sich wiederholendem Schriftverkehr in der Anwaltschaft genannt (Gebhard, 2022, S. 33). Bedenken äußerten die Teilnehmer:innen in Bezug auf die Frage, was das Basisdoku-

4. Wahrnehmung des Projekts Basisdokument

ment mehr könne als die genutzte Anwaltssoftware, in Bezug auf die Akzeptanz des Vorschlags bei älteren Anwält:innen, hinsichtlich der Realitätsnähe und Praxistauglichkeit der vorgeschlagenen Strukturierung, sowie der Verhinderung einer anwaltlichen Taktik (S. 34). Als Faktoren, die die Akzeptanz in der Anwaltschaft beeinflussen, sahen die Teilnehmer:innen vor allem das Alter der Nutzer:innen, die Verpflichtung zur Nutzung und die Furcht vor Einschränkung der anwaltlichen Freiheit und taktischer Mittel (S. 35ff). Die gezeigten Prototypen wurden jedoch überwiegend als einfach zu bedienen und auch für die Anwaltschaft nützlich wahrgenommen (S. 38f). Insgesamt empfiehlt Gebhard, die Anwält:innen stärker in die Entwicklung des Basisdokuments einzubeziehen – dies sei auch von essenzieller Bedeutung für den Erfolg der Umsetzung in der Praxis (S. 45). Gleichzeitig sieht er die geringe Bereitschaft der Anwält:innen an der Teilnahme solcher Projekte als eine Hürde bei der Modernisierung des Zivilrechts (S. 46).

4.2. Wahrnehmung der Erprobung im Reallabor Strukturvorgaben

Neben dem Vorschlag des Basisdokuments hat auch die Erprobung im Reallabor Strukturvorgaben Aufmerksamkeit erfahren. Es wurden noch nicht viele Beiträge veröffentlicht, im Folgenden sollen drei davon knapp umrissen werden.

Bert (2022a) (Rechtsanwalt) stellt fest, dass es bei der Diskussion um die Modernisierung des Zivilprozess „an allen Ecken und Enden“ an Daten und Empirie fehlt. Deshalb sei es gut, wenn die Justiz experimentiere, ausprobiere und teste – wissenschaftlich begleitet und reflektiert. Die vielen verschiedenen Ideen seien zu einem kohärenten Ganzen zusammenzuführen: „Zu einem modernen Zivilprozess, der Technik intelligent nutzt, der aber nicht nur auf Beschleunigung zielt, sondern qualitativ hochwertige Prozesse in den Mittelpunkt stellt und Richterinnen und Richter in die Lage versetzt, richtige Entscheidungen zu treffen.“

In einem weiteren Beitrag kommentiert Bert (2022b) (Rechtsanwalt) die Erprobung des Basisdokuments im Reallabor detaillierter. Die allgemeine Skepsis gegenüber dem Vorschlag sieht er als kritischen Punkt bei der Gewinnung von Testwilligen und damit in der Problematik, „den Prototypen an den Testgerichten oft genug

4. Wahrnehmung des Projekts Basisdokument

zu erproben, um valide Aussagen ableiten zu können“ (Abschnitt „Reallabor“). Er selbst teile die Skepsis gegenüber der Ersetzung des Tatbestands durch das Basisdokument und zweifle an der Sinnhaftigkeit der Strukturierung mittels des Basisdokuments in allen Verfahren. Er betont jedoch erneut, dass es der ZPO-Reform an Empirie fehle – das Reallabor biete eine Chance, das für das Basisdokument zu ändern.

Zwickel (2023) (Akademischer Oberrat an der FAU) schätzt, dass sich das Basisdokument im Reallabor (mit Ausnahme des Verbindlichwerdens des Urteilsbestands) „problemlos in das Spannungsfeld von Parteiherrschaft und Richtermacht“ einfüge (S. 91). Inhalt und Form des Vortrags blieben erhalten, nur der Modus sei verändert. Eine frühzeitige Erprobung des Basisdokuments hält Zwickel für unerlässlich.

4.3. Wahrnehmung des Basisdokumentprototyps

Zum Prototypen selbst scheint es in der juristischen Fachliteratur bisher keine Beiträge zu geben. Ich möchte jedoch an drei verschiedenen Beispielen zeigen, wie die Projektgruppe auch jetzt schon Rückmeldungen zum Prototyp erreichen. Dabei möchte ich berichten von (a) Rückmeldungen aus einem Workshop mit Richter:innen und einem Workshop mit Anwäl:innen, (b) einem LinkedIn-Post eines Anwalts, der den Prototyp aus Interesse ausprobiert hat und (c) meinem Besuch bei der Geschäftsstelle einer Zivilkammer in Regensburg zur Demonstration von eIP (elektronisches Integrationsportal) am 16.03.2023.

In den Workshops der Projektgruppe im November 2022 wurden den teilnehmenden Richter:innen und Anwäl:innen das Forschungsprojekt und der Prototyp vorgestellt und darüber diskutiert. In beiden Workshops wurde ein Protokoll erstellt, das mir als Grundlage für die Auswertung dient. Am Workshop für die Anwäl:innen nahmen 18 Rechtsanwält:innen teil (5 weiblich, 13 männlich). Am Workshop für die Richter:innen nahmen 29 Richter:innen teil (12 weiblich, 17 männlich). Die Projektgruppe war (zusätzlich) mit 13 bzw. 14 Teilnehmenden vertreten. Von den anwesenden Rechtsanwält:innen wurden sowohl der Prototyp als auch

4. Wahrnehmung des Projekts Basisdokument

der Vorschlag des Basisdokuments sehr kritisch kommentiert, ebenso wurden einige weitere Anforderungen an den Prototypen kommuniziert. Im Workshop der Richter:innen dagegen wurden Basisdokument und Prototyp kaum kommentiert, es ging stärker um weitere Anforderungen an den Prototypen.

Neben der allgemeinen Kritik am Vorschlag des Basisdokuments, wurden von Anwält:innen auch konkrete Kritikpunkte am Prototypen genannt: Etwa, dass der Prototyp sehr „mauslastig“ und damit nicht auf den Arbeitsablauf in den Kanzleien zugeschnitten sei und dass im Prototyp der Vortrag auf reinen Text reduziert werde. Weiterhin wurden sowohl von Anwält:innen, als von Richter:innen weitere Anforderungen an das digitale Basisdokument kommuniziert.

Ein weiteres spannendes Beispiel für Feedback zum Prototyp bildet ein LinkedIn-Post von Rechtsanwalt Dr. Matthias Friemelt.⁵⁴ Friemelt probierte den frei zugänglichen Prototypen aus, indem er einen fiktiven Fall konstruierte und einpflegte. Die Ergebnisse dieses Tests und seine Eindrücke vom Prototypen hat er auf LinkedIn gepostet. Er finde den Prototypen insgesamt gut nutzbar und glaube, dass das Projekt viel Potenzial habe. Einige Verbesserungsvorschläge bringt Friemelt auch direkt an: beispielsweise wünscht er sich mindestens zwei Gliederungsebenen mehr, hätte das Rubrum⁵⁵ lieber als Formular mit Eingabebeschränkungen und eine eigene Funktion für Beweisantritte.

Zuletzt möchte ich von meinem Besuch bei der Geschäftsstelle einer Zivilkammer in Regensburg zur Demonstration von eIP (elektronisches Integrationsportal) am 16.03.2023 berichten. Dabei stütze ich mich auf meine Feldnotizen, die im digitalen Anhang (Ordner /5_Feldnotizen) hinterlegt sind. Ausgehend vom Eingang eines Schriftsatzes beim Gericht erklärte die Mitarbeiterin die Arbeitsabläufe der Geschäftsstelle bzw. in eIP. Dabei wurde zum einen deutlich, dass der aktuelle Arbeitsablauf sich auf die Basisdokument-Dateien übertragen lässt, zum anderen aber auch, dass das Basisdokument-PDF wesentlich wichtiger ist als das Softwareent-

⁵⁴LinkedIn-Post zum Basisdokument, Matthias Friemelt: <https://www.linkedin.com/feed/update/urn:li:activity:7082614221124816896/>, abgerufen am 10.08.2023

⁵⁵Rubrum: Kurze, vorangestellte Zusammenfassung des Inhalts eines Schriftstücks. Im Prozessrecht sind im Rubrum formelle Inhalte (z.B. Anschriften der Prozessbeteiligten, Aktenzeichen u.ä.) des folgenden Schriftsatzes zusammengefasst.

4. Wahrnehmung des Projekts Basisdokument

wicklungsteam bisher angenommen hatte. Die Mitarbeiterin der Geschäftsstelle prüft z.B. bei jedem Satz, worum es geht und möchte deswegen auch im Basisdokument-PDF direkt erkennen können, welche Inhalte neu sind.

Diese drei Schilderungen sollen exemplarisch aufzeigen, auf wie unterschiedliche Arten die Projektgruppe (innerhalb und außerhalb der Erprobung im Reallabor Strukturvorgaben) Rückmeldungen zum Basisdokument(prototyp) bekommt. Diese Rückmeldungen fließen in die Arbeit der Projektgruppe und die Weiterentwicklung des Prototyps ein. Neue Anforderungen werden in der Projektgruppe diskutiert und für die Softwareentwicklung festgehalten.⁵⁶

⁵⁶Anforderungen werden als Issues im Repository gesammelt und abgearbeitet. Basisdokument-Repository: <https://github.com/UniRegensburg/basisdokument>, abgerufen am 29.11.2023

5. Projektstand im August 2023

Im Folgenden wird der Stand des Forschungsprojekts zu Beginn meiner Arbeit im August 2023 geschildert. Methodisch habe ich mithilfe einer formativen Evaluation des Forschungsprojekts nach den Richtlinien von Bergmann et al. (2005) einen Leitfaden für eine Fokusgruppe erstellt, die ich mit Projektbeteiligten aller Institutionen (Justizministerien und Lehrstühle für Verfahrensrecht und Medieninformatik) durchgeführt habe. So konnte ich weitere Einsicht in den aktuellen Stand des Projekts, die Projektziele und die Probleme im Reallabor Strukturvorgaben gewinnen und die Erwartungen der Projektbeteiligten an die Evaluation erfragen. Vor allem die Probleme, die im Reallabor bestehen, scheinen mir für die Eignungsanalyse einer Evaluationsmethode für die Erprobung wichtig: Sie bilden womöglich Einschränkungen und Rahmenbedingungen, die die Bewertung der Methode betreffen. Die folgenden Unterkapitel sind in diesem Sinne nicht nur als Projektbericht zu verstehen, sondern als tieferer Einblick in die Situation im Reallabor Strukturvorgaben, die es bei der Planung der Evaluation zu berücksichtigen gilt.

Zuletzt schildere ich einige neuere Entwicklungen im Projekt (Dezember 2023), die deutlich machen, wie die Projektgruppe auf Probleme reagiert hat und wie sich die Planung für das Projekt (auch ausblicksweise) weiterentwickelt hat.

5.1. Methode: Fokusgruppe mit Projektbeteiligten

Grundlage für die Planung der Fokusgruppendifkussion bildete eine formative Evaluation des Forschungsprojekts nach den Guidelines von Bergmann et al. (2005). Die Fragen aus dem Leitfaden habe ich mithilfe zur Verfügung stehender Informationen (Projekthomepage und Forschungsvertrag) und einer Besprechung mit Projektmitarbeitenden der Lehrstühle Medieninformatik und Verfahrensrecht am 18.07.2023 beantwortet. Diese Projektevaluation nach Bergmann et al. (2005) ist im

digitalen Anhang (Ordner /3_Fokusgruppe Projektstand) hinterlegt. Auf Basis dieser Vorarbeit habe ich den Leitfaden für eine Fokusgruppe am 04.08.2023 entworfen. Das Gespräch habe ich als Fokusgruppendifkussion durchgeführt, um auf einfache Weise ein breites Spektrum an Meinungen einzuholen (Lazar et al., 2017, S. 204). An der Fokusgruppe nahmen acht Mitglieder der Projektgruppe teil (drei weiblich, fünf männlich): zwei Beteiligte des Justizministeriums Bayern, eine Beteiligte des Justizministeriums Niedersachsen, zwei Beteiligte des Lehrstuhls für deutsches Verfahrensrecht und drei Beteiligte des Lehrstuhls für Medieninformatik. Mit Blick auf den zeitlichen Rahmen (eine Stunde) konnte ich nicht alle 56 Leitfragen der Guidelines mit den acht Teilnehmer:innen diskutieren. Der Fokus lag daher auf der Frage nach (a) dem konkreten Projektziel und seinen Erfolgskriterien, (b) den aktuellen Problemen mit Prototyp und Reallabor, (c) der geplanten Ergebnisoffenheit, (d) der gewählten Art der Praxiseinbindung (das Basisdokument an Gerichten in realen Verfahren einzusetzen) und ihrer Eignung zum Erreichen der Projektziele, (e) dem Verhältnis von Ergebnisumsetzung im konkreten Kontext zur Übertragbarkeit der Ergebnisse (Können mit dem konkreten Prototyp auch allgemein Einsichten zur digitalen Möglichkeit einer formellen Strukturierung des Parteivortrags gewonnen werden?) und (f) der Evaluation und den Erwartungen an sie.

Zur Auswertung der Gruppendiskussion wurde mithilfe der Aufnahme ein Protokoll erstellt, das den Teilnehmer:innen zur Kontrolle mit Möglichkeit zu Anmerkungen und Überarbeitungen zugesandt wurde.

Der Leitfaden für die Gruppendiskussion, das anonymisierte Protokoll und die Auswertung der Fokusgruppe sind im digitalen Anhang hinterlegt (Ordner /3_Fokusgruppe Projektstand).

5.2. Ergebnisse der Fokusgruppe

Dem Ansatz von Lazar et al. (2017, S. 311) folgend, habe ich die Inhalte des Protokolls in zwei Iterationen induktiv kodiert (Mayring, 2010). Da es mir darum geht, das Projekt inhaltlich zu verstehen, habe ich keine Quantifizierung der Nennungen und Kategorien angestrebt und vorgenommen. Die Ergebnisse der Auswer-

tung sind tabellarisch festgehalten und im digitalen Anhang hinterlegt (Ordner /3_Fokusgruppe Projektstand).

Die zentralen Einsichten zu Projektzielen, aktuellen Problemen und Evaluation werden im Folgenden zusammengefasst.

5.2.1. Projektziele

Zentrales Projektziel ist die Beantwortung der Frage, ob die Einführung des Basisdokuments als Werkzeug zur Strukturierung des Zivilprozesses lohnenswert ist oder nicht. Die Projektbeteiligten erhoffen sich ein tatsächliches Ergebnis in der Diskussion um die Strukturierung des Parteivortrags. In Bezug auf das Basisdokument möchte man Erkenntnisse darüber gewinnen, ob diese Methode der Strukturierung funktioniert (und ob die Anwält:innen das Angebot zur Strukturierung nutzen), ob das den Verfahrensbeteiligten tatsächlich einen Vorteil brächte und ob man durch den Praxistest Bedenken der Anwält:innen beseitigen könnte. Daneben sollen auch Erkenntnisse zur Nutzbarkeit des eingesetzten Basisdokumentprototyps gewonnen werden.

5.2.2. Aktuelle Probleme im Reallabor

Die aktuellen Probleme im Forschungsprojekt verteilen sich auf sechs Kategorien, die zum Teil nicht ganz trennscharf sind und im Folgenden anhand der Aussagen der Beteiligten dargelegt werden. Ein großes Problem im Reallabor Strukturvorgaben ist der *rechtliche Rahmen* bzw. die Regelungen des elektronischen Parteiverkehrs: Das Reallabor Strukturvorgaben ist an diese Regelungen gebunden, sodass die Basisdokument-Dateien kompliziert ausgetauscht werden muss – außerhalb des Prototyps. Dadurch müssen die unterschiedlichen Versionen der Basisdokument-Dateien händisch verwaltet werden. Durch diese Rahmenbedingungen kam es bisher auch zu den meisten *Problemen der Nutzer:innen*. Viele Nutzer:innen sind am Anfang überfordert von dem Prozess, zum Teil fehlt auch einfach das Wissen, wie das Basisdokument innerhalb der eigenen Systeme genutzt werden kann oder muss. Das bedingt dann auch einige Probleme in der *Kommunikation*: Der Verfahrensab-

lauf für den Austausch der Basisdokument-Dateien entspricht nicht den Vorstellungen der Projektbeteiligten für eine spätere Umsetzung – im aktuellen Reallabor Strukturvorgaben ist eine andere Umsetzung wegen der rechtlichen Rahmenbedingungen aber nicht möglich. Trotz dieser Probleme wird das umfangreiche Supportangebot von den Teilnehmer:innen nicht sehr häufig genutzt. Gründe, warum sich jemand gegen die Nutzung des Basisdokumentprototyps entschieden hat, werden kaum kommuniziert. Problematisch sehen die Teilnehmer:innen auch die zur Verfügung stehenden *Ressourcen*: sowohl zeitlich (es schien zu langwierig, eine Experimentierklausel für das Projekt zu bekommen), als auch personell (Personal für Informationsmanagement und Kommunikation ist knapp). Zusätzlich bremsen Personalwechsel bei den Gerichten den Projektverlauf. Auch bei den Tester:innen im Reallabor Strukturvorgaben sind die Ressourcen knapp – alle haben so viel zu tun, dass sie einfach keine Kapazitäten mehr haben, sich auch noch mit dem Basisdokument und dem Prototyp zu beschäftigen. Hätte man mehr Zeit und Möglichkeiten, mit den Testwilligen zu sprechen und das Projekt zu erklären, könnten viele Probleme vermieden und mehr Begeisterung für das Projekt geweckt werden. Weitere *Testwillige zu gewinnen* stellt ein weiteres Problem für das Reallabor Strukturvorgaben dar. Für die Anwerbung sei der aktuell komplizierte Arbeitsprozess mit den Basisdokument-Dateien ein Nachteil, ebenso bestehen weiterhin Bedenken der Anwaltschaft bezüglich potenzieller Einschränkungen des Vortrags und Verlagerung von Aufgaben des Gerichts auf die Anwält:innen. Auch für die *Weiterentwicklung der Software* bedeutet die Erprobung im Reallabor Strukturvorgaben aktuell einige Herausforderungen. Für die Entwicklung sei es schwierig, sich auf die Aussagen einzelner Personen verlassen zu müssen. Denn so sei unklar, ob eine Anforderung eine Einzelmeinung ist oder ob sie für alle Nutzer:innen relevant ist. Auch die Priorisierung einzelner Funktionen ist durch diesen Umstand erschwert.

5.2.3. Erwartungen an die Evaluation

Das zivilprozessuale Erkenntnisinteresse als Projektziel spiegelt sich erwartungsgemäß auch in den Erwartungen an die Evaluation und den Abschlussbericht wider.

Man wolle Rückmeldungen, Erkenntnisse und Interessen aller betroffenen Berufsgruppen einfangen. Es sei ein breiter Ansatz für die Evaluation notwendig (so soll z.B. auch das Strukturierungsgespräch Teil der Evaluation sein) und die angestrebte Ergebnisoffenheit solle auch im Abschlussbericht reflektiert werden. Bezogen auf das Basisdokument erhofft man sich Einsichten dazu, ob der Prozess für alle Beteiligten übersichtlicher, transparenter und damit effizienter wird. Ebenso soll evaluiert werden, was am Basisdokument gut bzw. schlecht umgesetzt ist. Dabei müsse klar sein, dass man nur den eingesetzten Prototypen evaluieren kann – anhand dessen soll auch identifiziert werden, was technisch gut funktioniert, sodass entsprechende Empfehlungen gegeben werden können. Außerdem soll die Evaluation auch darstellen, wie sich die Software weiterentwickelt hat und welche Faktoren zu dieser Weiterentwicklung geführt haben.

5.3. Neuere Entwicklungen im Projekt (Dezember 2023)

Im Rahmen einer Informationsveranstaltung bei der Rechtsanwaltskammer (RAK) München am 05.12.2023⁵⁷ konnte ich beobachten, wie die Projektgruppe auf Teile der oben geschilderten Probleme reagiert hat und wie das Projekt inzwischen von Anwält:innen wahrgenommen wird.⁵⁸ Meine Feldnotizen sind im digitalen Anhang hinterlegt (Ordner /5_Feldnotizen).

Vor allem fällt auf, dass das Projekt inzwischen unter dem Titel „Reallabor Basisdokument“ läuft – ich werde in dieser Arbeit aus Konsistenzgründen aber weiterhin den Begriff „Reallabor Strukturvorgaben“ nutzen. Die Kommunikation der Projektgruppe hat sich stark von dem Begriff der „Strukturvorgabe“ gelöst – und räumt damit einige Bedenken der Anwält:innen direkt aus. Auch in der Demonstration eines Musterprozesses (mit mehreren Wechseln zwischen Anwalt der Klagepartei, Richter am Gericht und Anwältin der Beklagtenpartei) weisen die Projektmitglieder mehrfach darauf hin, dass keine Strukturierung vorgegeben wird – weder vom Ge-

⁵⁷RAK München, „Infoveranstaltung zum Reallabor Basisdokument“: <https://seminare.rak-muenchen.de/64008-14-infoveranstaltung-zum-reallabor-basisdokument-2207974/>, abgerufen am 11.12.2023

⁵⁸LinkedIn-Post der RAK zur gelungenen Veranstaltung: https://de.linkedin.com/posts/rak-munich_rakm%C3%BCnchen-activity-7141411279507767297-mCwf, abgerufen am 18.12.2023

richt noch von der Klagepartei, die den ersten Schriftsatz in Form von Beiträgen im Basisdokument verfasst. Der Parteivortrag im Basisdokument(-prototyp) bildet eine gemeinsame Sammlung von Prozessstoff und bietet für alle Beteiligten jederzeit präsent eine Gesamtübersicht. Die Projektgruppe demonstriert und kommentiert das ausführlich. Die anwesenden Anwält:innen äußern diesbezüglich in der Frage- und Diskussionsrunde keine Bedenken, die Strategie der Projektgruppe scheint also erfolgreich zu sein. Ein neues *Framing* des Projekts führt in der Kommunikation potenziell wohl zu anderen Problemen (z.B. begriffliche Verwirrung), aber die Entscheidung der Projektgruppe, die Änderung vorzunehmen, um die Sorgen der Anwält:innen abzuschwächen, scheint nachvollziehbar. Ob das aber auch zu einer höheren Beteiligung an der Erprobung führt, bleibt abzuwarten. Die Anzahl der testwilligen Anwält:innen ist weiterhin eher gering: von Richter:innen wurden bis Dezember 2023 etwa 50 Verfahren gemeldet (teils bei vielversprechenden Verfahren oder bei Anregung der Nutzung des Basisdokuments, teils erst bei Vorliegen der Klage im Basisdokument), *Matches* (beide Parteien sind zur Teilnahme bereit) sind aber deutlich weniger zustande gekommen (niedriger einstelliger Bereich).

Auf Nachfrage geben die Projektmitglieder auch einen Ausblick zur weiteren Projektperspektive: Für ein anderes Modernisierungsprojekt (formularbasierte Prozesse) ist gerade ein Antrag mit Experimentierklausel in Arbeit, der eine Justizplattform erwähnt. Diese Plattform ist dem ähnlich, was für den Dateiaustausch des Basisdokuments auch sinnvoll scheint. Die Hoffnung der Projektgruppe ist, dass man darauf aufbauen oder sich für das Basisdokument auf ein bewährtes System (Justizplattform) stützen kann. Perspektivisch wünscht sich die Projektgruppe eine weitere Erprobung des Basisdokuments im Zivilprozess mit Experimentierklausel (also ohne Bindung an die ERVV). Der Basisdokumentprototyp soll dafür noch weiterentwickelt werden, um z.B. den Dateiaustausch zu integrieren.

6. Forschungsmethoden der Mensch-Maschine-Interaktion

Im Folgenden werden der Ansatz des *User Centered Design* und die gängigen Forschungsmethoden der Mensch-Maschine-Interaktion nach Lazar et al. (2017) vorgestellt. Diese Vorstellung soll die Basis für die Analyse und Bewertung von Evaluationsmethoden der Mensch-Maschine-Interaktion zur Anwendung in Reallaboren bilden. Ein Grundverständnis der Methodik ist zentral, um die Eignung einer Evaluationsmethode für das Reallabor Strukturvorgaben beurteilen zu können.

Methodisch soll der Ansatz des *User Centered Design*-Prozesses (UCD) den Rahmen des Forschungsprojekts bilden (Forschungsvertrag, 2022, S. 2). Im Folgenden wird dieser Ansatz in Absatz 6.1 vorgestellt. Das *User Centered Design* beschreibt einen Prozess, macht selbst jedoch keine Aussagen zu konkreter Methodik. Die Kapitel 6.2 bis 6.8 sollen, der Taxonomie von Lazar et al. (2017) folgend, einen Überblick über die gängigen Forschungsmethoden der HCI geben. Ich habe mich für die Taxonomie von Lazar et al. (2017) entschieden, weil sie eine recht offene zu sein scheint: Einzelne Methoden werden vorgestellt, ihre Vor- und Nachteile und ihre Einsatzmöglichkeiten benannt. Andere Autor:innen (z.B. Butz & Krüger (2017)) klassifizieren die einzelnen Methoden stärker nach ihrem Einsatzgebiet und unterscheiden dabei z.B. Anforderungserhebung und Evaluation. Diese Festlegung der Methoden auf einzelne Einsatzmöglichkeiten scheint mir dabei nicht sinnvoll umsetzbar, denn die meisten der vorgestellten Methoden lassen sich zu verschiedenen Zeitpunkten im UCD-Prozess und für verschiedene Ziele einsetzen. In der HCI ist es außerdem üblich, Methoden im Laufe einer Systementwicklung oder -evaluation sinnvoll zu kombinieren (Lazar et al., 2017, S. 130, S. 150, S. 183, S. 224).

Nach dem Ansatz des *User Centered Design* werden im Folgenden die einzelnen Methoden vorgestellt.

6.1. User Centered Design

User Centered Design ist ein Ansatz für einen Prozess zum Entwurf von Nutzendenschnittstellen (Butz & Krüger, 2017, S. 111). Dieser Entwicklungsansatz „integriert die Benutzbarkeit als gleichwertiges, wenn nicht sogar als entscheidendes Element in den Design- und Entwicklungsprozess“ (S. 111). Die Benutzbarkeit (*Usability*) wird definiert als „Ausmaß, in dem ein System, ein Produkt oder eine Dienstleistung durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um festgelegte Ziele effektiv, effizient und zufriedenstellend zu erreichen“ (DIN EN ISO 9241-210, 2011, S. 7). Oberstes Ziel des UCD-Prozesses ist die frühe und umfassende Einbindung der zukünftigen Nutzer:innen (Butz & Krüger, 2017, S. 112). Die vier Hauptelemente des UCD nach Butz & Krüger (2017, S. 112) sind das *Verstehen* der technischen und menschlichen Beschränkungen und Anforderungen, das *Designen* eines abstrakten, konzeptuellen und eines konkreten, physischen Vorschlags, das *Vergegenwärtigen* von Zwischenlösungen oder Designalternativen und das *Evaluieren* all dieser Schritte. Der ganze Prozess ist iterativ angelegt: Die Überarbeitung des Designs findet kleinteilig und jeweils gefolgt von einem Evaluationsschritt statt (S. 115). Am Ende des Designprozesses steht die Umsetzung des Designvorschlags aus der letzten Iteration (S. 115).

Dieser Ansatz der mensch-zentrierten Gestaltung ist in der Norm DIN EN ISO 9241-210 (2011) standardisiert. In Abbildung 5 ist der Prozess dargestellt.

Methoden zur Bewertung der Benutzerfreundlichkeit wurden in den 1990er Jahren aktiv entwickelt und sind der Standardansatz für *User Centered Design* (Salovaara et al., 2017, S. 4).

Wie in Kapitel 2 beschrieben, folgt die Entwicklung der Basisdokumentprototypen am Lehrstuhl Medieninformatik seit 2021 diesem Ansatz.

6.2. Kontrollierte Experimente

Experimente werden in der HCI häufig eingesetzt, um Benutzermodelle oder Aufgabenmodelle zu entwickeln und zu ändern, verschiedene Designlösungen zu be-

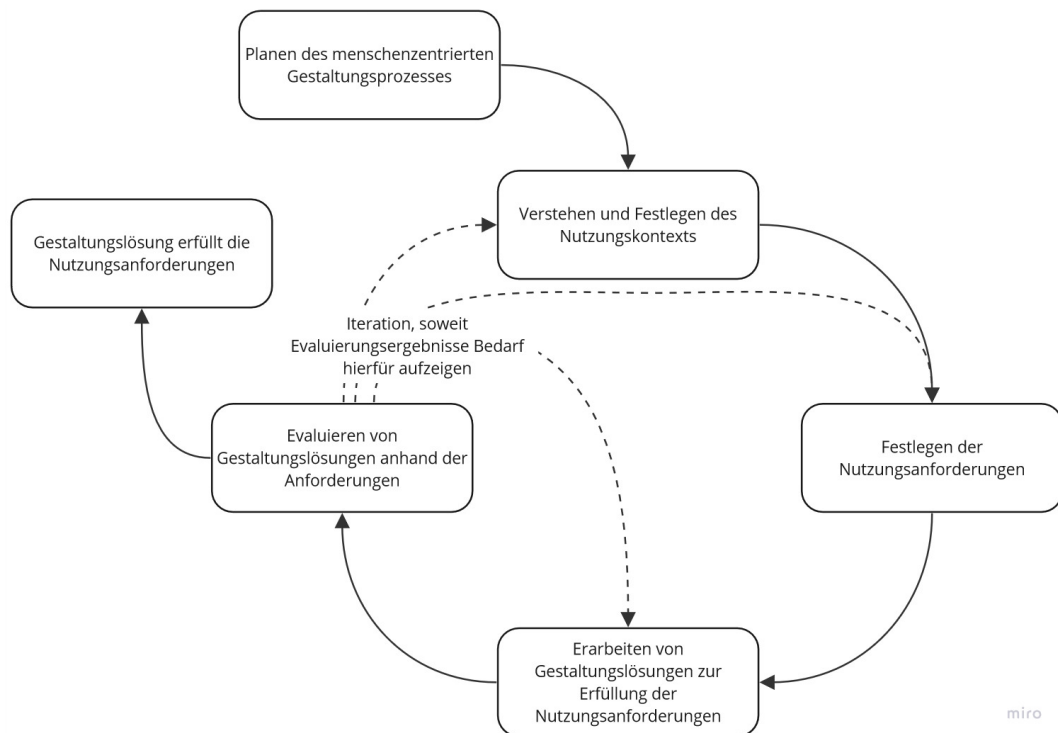


Abbildung 5: Wechselseitige Abhängigkeit menschenzentrierter Gestaltungsaktivitäten im *User Centered Design*-Prozess (DIN EN ISO 9241-210, 2011, S. 15).

werten und verschiedene andere kritische Fragen zu beantworten, z. B. zur Technologieakzeptanz (Lazar et al., 2017, S. 45). Die größte Stärke experimenteller Forschung ist die Möglichkeit, kausale Beziehungen zu identifizieren (S. 41). Experimente basieren auf einer oder mehreren testbaren Forschungshypothesen und zielen darauf ab, sie zu validieren (S. 45). Typischerweise werden mindestens zwei Bedingungen oder Gruppen getestet, abhängige Variablen werden quantitativ gemessen und die Ergebnisse werden durch statistische Signifikanztests analysiert (S. 45). „Kontrolliert“ bezieht sich dabei auf die Rahmenbedingungen des Experiments: Alle Variablen werden überwacht, das Experiment meist in einem Labor durchgeführt.

6.3. Umfragen

Eine Umfrage (engl. *survey*) ist ein klar definierter und gut formulierter Satz von Fragen, auf die eine Person antworten soll (Lazar et al., 2017, S. 105). Typischerweise füllt die Person die Umfrage (auf Papier oder online) aus, ohne dass die Anwesenheit eines Forschenden notwendig ist – die erhobenen Daten sind deswegen nicht so

tiefgehend wie bei anderen Methoden (S. 105). Die Stärke der Methode liegt darin, eine große Anzahl an Antworten schnell von einer (womöglich geografisch verteilten) Gruppe von Personen zu bekommen (S. 105). Umfragen können genutzt werden, um einen Eindruck davon zu bekommen, wie Personen mit einer Technologie interagieren, welche Probleme auftauchen und welche Aktionen sie durchführen (S. 105). Sie sind gut geeignet, um Daten zu Einstellungen, Bekanntheitsgrad, Ziele, Rückmeldungen über Nutzungserfahrungen, Merkmale der Nutzer:innen und Vergleiche im Zeitverlauf zu erheben (S. 105). Man kann Umfragen so aufbauen, dass die Ergebnisse statistisch verwertbar sind (S. 105). Die einzelnen Fragen müssen klar, eindeutig und unvoreingenommen sein (S. 130).

6.4. Tagebuchstudien

Ein Tagebuch (engl. *diary*) ist ein Dokument, das von einer Person erstellt wird, die regelmäßig Aufzeichnungen über Ereignisse in ihrem Leben macht, und zwar zu dem Zeitpunkt, an dem diese Ereignisse eintreten (Lazar et al., 2017, S. 135). In der HCI werden Tagebücher als Forschungsinstrument genutzt, um z.B. Nutzungsmuster neuer Technologien zu verstehen (S. 150). Tagebuchstudien sind nützlich, wenn Technologie unterwegs genutzt wird (S. 150) und für Informationen, die sich verändern (wie z.B. die Stimmung) oder Ereignisse, die mehrfach oder regelmäßig auftreten (S. 135). Tagebuchstudien sind sinnvoll einsetzbar, um über einen längeren Zeitraum Daten zu sammeln.

6.5. Fallstudien

Eine Fallstudie (engl. *case study*) ist eine eingehende Untersuchung eines bestimmten Falles (oder einer kleinen Anzahl von Fällen) in einem bestimmten realen Kontext (Lazar et al., 2017, S. 153). Sie kann eine nützliche Methode sein, um Anforderungen zu sammeln oder Interfaces zu evaluieren – besonders, wenn es schwierig oder unmöglich ist, eine größere Stichprobe zu akquirieren (S. 153). Die vier Kernaspekte sind (a) die eingehende Untersuchung einer kleinen Anzahl von Fällen, (b) die Untersuchung im realen Kontext, (c) die Nutzung mehrerer Datenquellen

und (d) der Schwerpunkt auf qualitativen Daten und deren Analyse (S. 156). Die Ziele von Fallstudien fallen meist in die Kategorien Untersuchung (engl. *exploration*: neuartige Probleme oder Situationen verstehen), Erklärung (engl. *explanation*: Modelle zum Verstehen eines Kontexts von Technologienutzung entwickeln), Beschreibung (engl. *description*: ein System, einen Kontext von Technologienutzung oder einen Designprozess dokumentieren) und Demonstration (engl. *demonstration*: zeigen, wie ein neu eingesetztes Tool erfolgreich genutzt wird) (S. 159). Die kleine Stichprobe einer Fallstudie ist ein Nachteil für ihre Validität und Reliabilität, Ergebnisse sollten also entsprechend interpretiert und kommuniziert werden (S. 183). Dennoch können Fallstudien unser Verständnis komplexer Phänomene vertiefen – sie können dabei helfen, Designdetails zu identifizieren und konkrete Darstellungen von Bedürfnissen, Motivationen, Erfolgen und Misserfolgen zu liefern (S. 183).

6.6. Interviews und Fokusgruppen

Ein ebenfalls schmalere Ansatz, der zu tiefergehenden Ergebnissen führt (verglichen mit Umfragen), sind Befragungen (S. 187). Direkte Diskussionen mit betroffenen Teilnehmer:innen werden entweder als Interviews mit einzelnen Personen oder Fokusgruppen mit mehreren Nutzer:innen gleichzeitig durchgeführt (S. 187) – die Entscheidung wird von Kompromissen hinsichtlich Zeit, Zweckmäßigkeit, Tiefe und Schwierigkeit beeinflusst (S. 224). Bei diesen Befragungen kann es sich um ausführliche Gespräche handeln, die darauf abzielen, Modelle zu erstellen, um zu erklären, wie Systeme verwendet werden und warum, sie können dazu beitragen, ein Verständnis für Bedürfnisse und Reaktionen auf Schnittstellen zu entwickeln oder Erkenntnisse darüber zu gewinnen, warum der letzte Entwurf gescheitert ist (S. 224). Die Befragungen können unterschiedlich stark strukturiert sein (S. 198) und können bei Bedarf in einem spezifischen Kontext stattfinden, sodass Teilnehmer:innen Dinge auch demonstrieren können, statt nur zu erklären (S. 200) – bis hin zur *Contextual Inquiry* (S. 201) oder der Nutzung von (*Technology*) *Probes*, um Feedback oder Reaktionen hervorzurufen (S. 203).

6.7. Ethnographie

Die Ethnographie basiert auf der Vorstellung, dass ein wirkliches Verständnis komplexer menschlicher Praktiken und Kontexte ein tiefgreifendes, engagiertes Studium und somit zumindest ein gewisses Maß an Beobachtung und Beteiligung erfordert (Lazar et al., 2017, S. 231). Als beschreibende Technik arbeitet die Ethnographie meist induktiv und bewegt sich von Rohdaten zur Identifizierung von Mustern, die regelmäßig in den Daten auftreten, und zu allgemeinen Theorien, die die Muster erklären (S. 231f). In der HCI wird Ethnographie oft als erster Schritt verwendet, um eine Gruppe von Nutzer:innen, ihre Probleme, Herausforderungen, Normen und Prozesse zu verstehen, mit dem Ziel, eine Technologie für sie oder mit ihnen zu entwickeln (S. 256). Wie Fallstudien setzen auch ethnographische Studien auf mehrere Datenerfassungstechniken, um eine umfassende Perspektive zu gewinnen (S. 243). Der Unterschied zu Fallstudien liegt in der Nutzung von Theorien: bei Fallstudien leiten häufig (Hypo)-Thesen die zentrale Fragestellung, während die Ethnographie induktive Strategien nutzt (S. 232). Außerdem setzt die Ethnographie auf einen längeren Zeitraum der Beobachtung als bei anderen Methoden üblich (S. 232).

6.8. Usability Testing

Das grundlegende Ziel von *Usability Testing* ist es, die Qualität eines Interfaces zu verbessern, indem durch den Test problematische Bereiche oder Mängel im Interface gefunden werden (Lazar et al., 2017, S. 264). Auch was im Interface gut funktioniert, ist ein interessantes Ergebnis – denn das soll beibehalten werden (S. 264). Methodisch nutzt *Usability Testing* fast immer eine Kombination der bereits vorgestellten Methoden der HCI (S. 265), am häufigsten Methoden des *Experimental Design* und der Ethnographie (S. 266). So werden z.B. beim *Participatory Design* ethnographische Methoden als Teil einer Systementwicklungsmethode genutzt (S. 266). Die Details des Usability-Tests (wie das Entwicklungsstadium des Prototyps, Ort der Testung, Grad der Formalität, Aufgabenliste, Anzahl der Teilnehmer und verwendete Metriken) werden durch Budget, Zeitplan und Organisation des Projekts

bestimmt (S. 295).

Usability-Tests lassen sich in drei Kategorien unterscheiden: expertenbasiertes Testen, automatisiertes Testen und nutzerbasiertes Testen (S. 267). In der Praxis wird häufig eine Kombination verschiedener Methoden genutzt.

6.8.1. Methoden des expert:innenbasierten Testens

Diese Testart besteht aus strukturierten Inspektionen des zu untersuchenden Interfaces, auch bezeichnet als *Expert Review* (Lazar et al., 2017, S. 268). Diese analytische Methode wird typischerweise vor Tests mit Nutzer:innen angewandt, um einige der offensichtlichsten Schnittstellenmängel aufzudecken (z. B. verwirrende Formulierungen, inkonsistentes oder fehlendes Layout und Farbinconsistenz), sodass diese Art von Mängeln die Nutzer:innen im Test nicht ablenkt (S. 268). Die gebräuchlichsten *Expert Reviews* sind heuristische Bewertungen, Konsistenzprüfung und *cognitive walkthrough* (S. 268).

Heuristische Bewertung. Ein:e Expert:in analysiert das Interface anhand einer Reihe von Heuristiken (S. 268). Es gibt viele verschiedene Heuristiken für diese Art von *Usability Testing*, die bekanntesten sind vermutlich die acht goldenen Regeln von Shneiderman et al. (2017) (S. 269). Ebenfalls häufig genutzt sind die zehn Heuristiken von Nielsen (1994) (Butz & Krüger, 2017, S. 140ff).

Konsistenzprüfung. Ein:e oder mehrere Expert:innen überprüfen eine Reihe von Screens oder Webseiten auf Konsistenz in Bezug auf Layout, Farbe, Terminologie oder Sprache – zum Teil anhand spezifischer Styleguides (z.B. *Corporate Design*) (Lazar et al., 2017, S. 269).

Cognitive Walkthrough. Expert:innen simulieren, wie Nutzer:innen eine Reihe von Aufgaben „durchgehen“ (Lazar et al., 2017, S. 269). Diese Methode ist eher explorativ und nicht regelbasiert (S. 269).

Weniger verbreitet ist die Durchführung von *Guidelines Reviews*, in denen Expert:innen Interfaces mit einem großen Set (10-200) von vorher festgelegten Interface-Guidelines abgleichen (S. 269). Beispiele für solche Guidelines sind die *Web Content*

Accessibility Guidelines (WCAG) des W3C⁵⁹ oder Richtlinien für Interfacegestaltung von z.B. Microsoft⁶⁰, Apple⁶¹, KDE⁶² oder GNOME⁶³.

6.8.2. Automatisiertes Testen

Beim automatisierten Testen analysiert eine Software anhand festgelegter Guidelines das Interface (Lazar et al., 2017, S. 270). Diese Guidelines können z.B. die gleichen sein, die für *Guideline Reviews* von Expert:innen genutzt werden. Es steht eine große Menge an Tools für automatisiertes Testen zur Verfügung – für Aspekte der *Accessibility* z.B. Anwendungen wie Deque WorldSpace⁶⁴ oder web-basierte Tools wie A-Checker⁶⁵, WAVE⁶⁶, Functional Accessibility Evaluator⁶⁷ (S. 271). Andere Aspekte können z.B. Performanz, Suchmaschinenoptimierung (SEO), Sicherheit, *Usability*, Struktur oder *Experience* sein (Namoun et al., 2021).

6.8.3. Methoden des nutzer:innenbasierten Testens

Bei dieser Art des Testens führen repräsentative Nutzer:innen meist repräsentative Aufgaben im Interface aus (Lazar et al., 2017, S. 271). Nutzer:innenbasierte Tests bilden den Schwerpunkt der Usability-Evaluation (S. 268), sie werden in formative (während des Designprozesses), summative (wenn die großen Designentscheidungen schon getroffen wurden) und Validationstests (kurz vor Release wird das neue Interface mit Benchmarks für andere Interfaces verglichen) unterschieden (S. 271f). Verschiedene Autor:innen nutzen verschiedene Definitionen der Begrifflichkeiten, der gemeinsame Kern ist aber, dass frühere, formative Usability-Tests sich mehr auf qualitatives Feedback, Beobachtung durch Moderator:innen und Problemfindung

⁵⁹WCAG des W3C: <https://www.w3.org/WAI/standards-guidelines/wcag/>, abgerufen am 28.08.2023

⁶⁰Interfacerichtlinien von Microsoft: <https://learn.microsoft.com/en-us/windows/win32/appuistart/-user-interface-principles>, abgerufen am 28.08.2023

⁶¹Interfacerichtlinien von Apple: <https://developer.apple.com/design/human-interface-guidelines/>, abgerufen am 28.08.2023

⁶²Interfacerichtlinien von KDE: <https://develop.kde.org/hig/>, abgerufen am 28.08.2023

⁶³Interfacerichtlinien von GNOME: <https://developer.gnome.org/hig/>, abgerufen am 28.08.2023

⁶⁴Deque WorldSpace: <https://www.deque.com/>, abgerufen am 25.08.2023

⁶⁵A-Checker: <https://achecks.org/achecker/>, abgerufen am 25.08.2023

⁶⁶WAVE: <https://wave.webaim.org/>, abgerufen am 25.08.2023

⁶⁷Functional Accessibility Evaluator: <https://fae.disability.illinois.edu/>, abgerufen am 25.08.2023

konzentrieren, während summative Usability-Tests sich mehr auf aufgabenbezogene Messungen, Metriken und quantitative Messmethoden fokussieren (S. 273).

Die gebräuchlichsten quantitativen Messungen sind Aufgabenperformanz, Zeitperformanz und Zufriedenheit der Nutzer:innen (S. 288). Weitere Metriken sind meist auf den Kontext der Aufgabe abgestimmt, z.B. die Anzahl der Fehler, die durchschnittliche Zeit einen Fehler zu korrigieren, wie oft die Suche verwendet wurde oder wie oft geklickt wurde (S. 288). Quantitative Daten werden häufig statistisch ausgewertet, um z.B. Nutzendengruppen oder Interfaceversionen zu vergleichen. Als qualitative Methode wird oft *Thinking Aloud* eingesetzt: dabei werden Teilnehmer:innen dazu ermuntert, beim Lösen der Aufgaben „laut mitzudenken“ (S. 288). Zur Messung der Zufriedenheit und anderer Metriken werden häufig standardisierte Fragebögen genutzt (S. 288). Eine Sammlung solcher Fragebögen ist im Anhang in Kapitel A.1 als Tabelle zusammengestellt. Quellen für diese Sammlung sind sowohl Lehrbücher (Chesnut & Nichols (2014); Lazar et al. (2017); Shneiderman et al. (2017); Butz & Krüger (2017), als auch Sammlungen von UX-Methodensammlungen⁶⁸. In der Tabelle sind für jeden Fragebogen auch die Originalpublikationen hinterlegt. Diese Sammlung ist auf keinen Fall vollständig, kann aber einen Eindruck davon vermitteln, wie verschieden die Metriken sind, die operationalisiert werden. Die aufgeführten Fragebögen haben unterschiedliche Messungen zum Ziel, z.B. die Zufriedenheit (ASQ, QUIS), Erfahrung (UEQ, IMI) oder Akzeptanz (TAM) der Nutzer:innen, die Qualität (AttrakDiff, HED/UT), Konsistenz (ICTQ) oder Ästhetik (Aesthetics scale) der Anwendung, Aufgabenlast (NTLX), Affekt (PANAS, Affect Grid), Emotion (GAQ, SAM) oder *Usability* (CSUQ, ISOMetrics, PUTQ, SUMI, SUS, USE, WAMMI). Einen bibliografischen Überblick über 32 gängige Fragebögen im Bereich *Usability* und *User Experience* gibt die Arbeit von Apraiz Iriarte & Lasa Erle (2020). Manche dieser Fragebögen werden in Usability-Evaluationen im Anschluss an die Aufgabe ausgefüllt, manche im Anschluss an die Studie, manche eignen sich z.B. auch für die allgemeine Bewertung

⁶⁸UX-Methodensammlungen finden sich z.B. auf <https://experienceresearchsociety.org/ux/evaluation-methods/> oder <https://www.uxmethods.guru/catalog>, aufgerufen am 16.11.2023

nach Nutzung einer Website oder Anwendung (außerhalb einer Studie, im regulären Betrieb). Manche dieser Fragebögen werden immer wieder weiterentwickelt, manche auch angepasst an bestimmte Kontexte – wie z.B. der AttrakWork als Abwandlung des AttrakDiff für die Evaluierung der Nutzendenerfahrung von mobilen Systemen im Kontext des mobilen Nachrichtenjournalismus (Väättäjä et al., 2009). Mit der Entwicklung neuer Arten von Interfaces, Datenerhebung und Interaktionsmöglichkeit wie z.B. Computerspiele, *Eyetracking* oder *Virtual Reality* entstehen auch neue Methoden und Fragebögen (z.B. der *Game Experience Questionnaire* (IJsselstein et al., 2013)).

Die methodische Vielfalt der HCI-Forschung macht die Auswahl einer geeigneten Methode potenziell schwierig – bietet aber auch die Möglichkeit, Methoden passgenau zu wählen. Es gibt sehr viele (teils validierte) Fragebögen, viele selbstentwickelte Metriken werden genutzt (Vermeeren et al., 2010; Pettersson et al., 2018) und ein weites Angebot sich überschneidender Theorien und Methoden ist verfügbar. Die klassischen *Usability*-Studien haben für die Evaluation häufig mit kontrollierten Bedingungen gearbeitet – im Feld wurde eher beobachtet und Anforderungen erhoben. Das hat sich in den letzten Jahren allerdings verändert, sodass ich im folgenden Kapitel auf die Evaluation im Feld eingehe.

7. Evaluation im Feld

Im Folgenden wird die Evaluation im Feld in der Mensch-Maschine-Interaktion dargestellt. Nach einigen begrifflichen Einordnungen schildere ich die Rolle von Feldversuchen in der HCI, um dann auf die Evaluationsmethodik in der Feldforschung einzugehen. Ein Kernelement von Reallaboren ist der Einsatz einer innovativen Technologie im realen Kontext. Für die Planung der Evaluationsmethodik ist ein Blick in die Feldforschung also naheliegend.

Wenn man sich mit Evaluation im Feld (im Gegensatz zum Labor) beschäftigt, stößt man schnell auf eine Vielzahl von Begrifflichkeiten. Außer „im Feld“ (engl. „in the field“) sind auch die Begriffe „in situ“ und „in the wild“ oder Ausdrücke wie „situated interaction“ verbreitet.

Auch für den „Versuch“, die „Erprobung“, die „Studie“ sind verschiedene Bezeichnungen in Gebrauch – z.B. „trial“ oder „field experiments, deployments, evaluations, field studies, technical probes“ (Brown et al., 2011, S. 1658). Brown et al. (2011) definieren aber eine Reihe gemeinsamer Funktionen dieser Experimente:

„Ein neues, in der Regel von den Forschenden entwickeltes System wird einer Reihe von Nutzenden zur Verfügung gestellt, die (oft implizit) aufgefordert werden, das System ‚natürlich‘ außerhalb des Labors zu nutzen, in dem es entwickelt wurde. Sie verwenden das System dann einige Stunden bis zu einem Jahr lang, häufig als Teil ihres täglichen Lebens, oder das System wird bei ihnen zu Hause oder an ihrem Arbeitsplatz eingesetzt. In eher experimentellen Versuchen kann das Verhalten der Nutzenden durch Aufgaben eingeschränkt werden, während andere Versuche auf solche Kontrollen verzichten und versuchen, eine ‚natürliche‘ Nutzung zu fördern. In einigen Fällen werden die Studien als Erweiterung von Experimenten durchgeführt, während in anderen

Fällen ein eher ethnographischer Analysemodus gewählt wird. Bei allen Versuchen werden Daten über die Nutzung des Systems erhoben“ (S. 1658)

Auch bei den Räumen der Erprobung herrscht große begriffliche Vielfalt, z.B. eben Reallabore, *Living Labs*, Innovationslabore, *Sandboxes*, Experimentierräume, Realexperimente, Pilotprojekte, ... – diese Räume bzw. Begrifflichkeiten werden natürlich nicht nur in der HCI genutzt, sondern lassen sich für verschiedene Fachbereiche sinnvoll einsetzen.

Das Ziel dieser Ansätze ist herauszufinden, „was passiert“ und Gestaltungsprinzipien oder Empfehlungen zu den Reaktionen der Nutzer:innen abzuleiten – meist durch eine Mischung aus qualitativen und quantitativen Aspekten (Brown et al., 2011, S. 1657).

7.1. Feldversuche in der HCI

Die ersten Feldversuche mit Technologien lassen sich Anfang der 1990er Jahre verorten, als man neben der Technologieerprobung im Labor auch eine breitere Untersuchung der Interaktion mit diesen Technologien anstrebte (Barkhuus & Rode, 2007; Brown et al., 2011; Rogers & Marshall, 2017). Diese frühen Feldversuche führten Forschende in ihrem eigenen (Arbeits)-Umfeld aus, z.B. *Portholes* im EuroPARC (Dourish et al., 1996), das *Active Badge*-System (R. Harper et al., 1992) oder verschiedene Medienraumsysteme in den *Bell Labs* (Root, 1988).

Im Anschluss an diese ersten Versuche hat sich der Ansatz mehr und mehr zu einer zentralen Methode für die Untersuchung von Nutzenden-System-Interaktionen entwickelt (vgl. z.B. Benford et al. (1998); Smith et al. (2005); Chamberlain & Crabtree (2020)). Einen Schub hat das Feld auch durch die Forschung im Bereich *Ubiquitous Computing* und des *Internet of Things* erhalten (Chamberlain & Crabtree, 2020). Durch die enge Beziehung zur Umgebung bzw. dem Nutzungskontext und das Interesse an Systemen, die Nutzer:innen im Alltag unterstützen oder in den Alltag eingreifen, scheint es sinnvoll auf Feldversuche zu setzen, sodass sie sich inzwischen zu einer Standardmethode der HCI entwickelt haben (vgl. Brown et al. (2011, S.

1658) oder z.B. Bell et al. (2006); Benford et al. (2004); Rogers et al. (2007); Rogers & Marshall (2017); Chamberlain & Crabtree (2020)). Dabei wird auch verstärkt auf längere Erprobungszeiträume (mehrere Wochen oder Monate) gesetzt, da die technologischen Hürden für die Erprobung und das *Deployment* neuer Anwendungen und Hardware im Lauf der Zeit immer weiter sinken (Brown et al., 2011, S. 1658).

Eine weitere Entwicklung gab es durch die inhärenten Schwierigkeiten dieser Methode bzw. deren Einsatz – so wurden Aspekte z.B. des Heims und der Heimtechnologie wieder in das Labor zurückgebracht, in Form von *LivingLabs* wie dem *PlaceLab* des MIT (Intille et al. (2005), siehe auch Kapitel 3). Solche Räume ermöglichen es Forschenden, Unsicherheiten zu minimieren und dennoch viele Aspekte von Feldversuchen beizubehalten.

Der Feldversuch als Methode zur Untersuchung von Nutzendeninteraktionen wurde und wird durchaus kontrovers gesehen und diskutiert. Ein prominenter Vorwurf ist, dass Feldversuche bzw. die neue Technologie existierende Praktiken stören und Forschende stattdessen Praktiken unterstützen, die sich noch nicht herausgebildet haben (Crabtree, 2004) – so würde z.B. die Einführung von Technologie in den Haushalt das „alltägliche“ häusliche Leben unterbrechen und damit die Möglichkeit einschränken, das Ziel zu erreichen, das „Alltägliche“ zu verstehen (Tolmie & Crabtree, 2008). In einigen Diskussionen wurde der Nutzen der Praxis selbst und ihr Beitrag zur Forschung infrage gestellt. Davies (2005) argumentiert, dass Feldforschung nicht nur zeitaufwendig, sondern oft auch das falsche Werkzeug ist:

„Das Problem ist oft, dass es kein eigentliches Konzept gibt, das bewiesen werden muss. Entweder wurde das Konzept bereits als praktikabel erwiesen (es muss wirklich nicht noch einmal bewiesen werden, dass wir einen kontextabhängigen Reiseführer bauen können), es wird nie angezweifelt (wir wissen, dass wir standortbezogene Dienste bauen können) oder es wird vom Demonstrator nicht wirklich bewiesen.“
(Davies, 2005, S. 2)

Feldforschung ist nicht nur zeitaufwendig, sondern auch arbeitsintensiv (für Forschende immer, für Beforschte teils auch) und dadurch teuer und mit der Investiti-

on von Ressourcen (finanziell und personell) verbunden (Chamberlain & Crabtree, 2020, S. 3). Weitere Einwände beziehen sich darauf, dass Feldversuche nur einen geringen Mehrwert für Studien mit Nutzer:innen bieten und vieles, was herausgefunden wird, einfacher in einer Laborstudie entdeckt würde (z.B. Kjeldskov et al. (2004)).

Dem entgegen zeigten Rogers et al. (2007), dass viele Probleme mit der Benutzerfreundlichkeit im Labor nicht auftreten, wenn man sie mit der Verwendung vor Ort vergleicht, und dass Feldversuche einen „kontextuellen Hintergrund“ bieten, „vor dem wir über die Gestaltung der Benutzererfahrung und des [Systems] nachdenken [können], und [... der] uns dafür [sensibilisiert], wie [unser System] in der Praxis genutzt werden würde (und nicht sollte).“ (S. 351).

Auch weiterhin wird dieser Diskurs in der Fachgemeinschaft geführt, z.B. bei einem Panel der *Conference on Human Factors in Computing Systems* (CHI) 2019 (C. Liu et al., 2019).

Besonders die Etablierung von *LivingLabs* und ähnlichen Einrichtungen führte außerdem zur Entwicklung von Begleitforschung (z.B. Schöpke et al. (2018); Wanner et al. (2018); Alavi et al. (2020); Tag et al. (2022); Anduschus et al. (2023)).

Feldversuche werden in der HCI inzwischen in vielen verschiedenen Bereichen eingesetzt, beispielhaft mit ebenso beispielhaften Nachweisen sind zu nennen: Gesundheit (Osman Andersen et al., 2021; Dürr et al., 2020; K. Koch et al., 2021), Informationspraktiken (Innocenti et al., 2022), kulturelles Erbe und Museumswesen (Galani & Kidd, 2019), Spiele (Benford et al., 2004; Bell et al., 2006), *Smart Home* (Tolmie & Crabtree, 2008), Nachhaltigkeit (Sjöman & Hesselgren, 2020), Robotik (Sabanovic et al., 2014) oder Usability-Evaluation von Software (Geszten et al., 2019). Seit einigen Jahren wird auch in der Feldforschung zum Teil auf künstliche Intelligenz gesetzt (z.B. bei Osman Andersen et al. (2021); Y. Liu et al. (2022)) und verschiedene Frameworks und Modelle wurden entwickelt (z.B. bei Ranasinghe et al. (2019); Youoku et al. (2020)).

7.2. Evaluationsmethodik in der Feldforschung

Im Rahmen der Feldforschung werden in der Regel verschiedene Methoden angewandt, darunter beispielsweise Beobachtung, Umfragen, Aufzeichnung der Nutzungsdaten und die Einbeziehung von Beteiligten in verschiedenen Kontexten durch Fokusgruppen, Co-Design-Sitzungen und Gruppensitzungen, um ihre Meinungen zu hören und sie ihre Bedenken äußern zu lassen (Rogers & Marshall, 2017, S. 7).

Die meisten Feldstudien wenden qualitative und ethnografische Methoden an, um die Nutzung neuer Technologien vor Ort zu untersuchen; die Stärke dieser Methoden liegt darin, dass sie es den Forschenden ermöglichen, die Auswirkungen der eingesetzten Technologien im Detail zu untersuchen und unerwartete Phänomene zu dokumentieren und zu erklären (Rogers & Marshall, 2017, S. 29). Der Ansatz der Evaluation im Feld strebt – als Gegensatz zu den HCI-typischen, kontrollierten Laborstudien – besonders auch nach höherer ökologischer Validität (siehe dazu z.B. Rogers & Marshall (2017, S. 79) oder Chamberlain & Crabtree (2020, S. 2)). Dies scheint vor allem auch mit Blick auf die Replikationskrise wichtig, die ausgehend von der Psychologie auch die HCI ergriffen hat und zu weiteren Diskussionen über die Validität von Forschungsergebnissen führt. Die Replikationskrise (*replication crisis*) bezieht sich auf die Unfähigkeit, wissenschaftliche Experimente bzw. deren Ergebnisse zu reproduzieren, und ist eine große Herausforderung für viele Bereiche der Wissenschaft. Eine ausführlichere Darstellung der Rolle der HCI in der Replikationskrise geben z.B. Echtler & Häußler (2018); Feger et al. (2019); Cockburn et al. (2020).

Bei der Evaluation geht es in der Forschung der letzten Jahre auch verstärkt um großflächigen Einsatz (*large scale deployment*, z.B. bei Babahmetović & Manresa-Yee (2019); Mamykina et al. (2021)) oder (semi)-automatische Evaluationsmethoden (z.B. bei Generosi et al. (2020); M. Koch et al. (2023)), die verstärkt auch auf quantitative Methoden setzen.

Die Empfehlungen für die Evaluation im Feld bleiben notwendigerweise eher abstrakt – zu vielfältig sind die Projekte, ihre Forschungsgebiete und die zu untersuchenden Technologien.

8. Evaluation im Reallabor Strukturvorgaben: Planung der Projektgruppe

Im Folgenden wird die bisherige Planung der Projektgruppe für die Evaluation des Reallabors Strukturvorgaben dargestellt.

Die Festlegung einer Evaluationsmethodik muss in deutschen Reallaboren nur für Projekte erfolgen, die eine Experimentierklausel nutzen. Im vorliegenden Reallabor Strukturvorgaben ist das nicht der Fall, sodass der geplante Testbetrieb unter den Regelbedingungen des Zivilprozesses (Althammer et al., 2023, S. 2) und die Ausgestaltung der Evaluierung innerhalb des Projekts erfolgen.

Erste Vorschläge und Ideen für die Evaluation aus der Projektgruppe wurden mit Beginn des Reallabors Strukturvorgaben veröffentlicht (Althammer et al., 2023). Hier wird auch verdeutlicht, dass die Evaluation zwei Perspektiven zum Ziel haben muss: zum einen wird es um die Evaluation der prototypischen Implementierung des digitalen Basisdokuments gehen, zum anderen um das übergeordnete regulatorische Erkenntnisinteresse mit Blick auf die Weiterentwicklung der Zivilprozessordnung in Hinblick auf Strukturvorgaben im Parteivortrag. Aus Sicht des *software* bzw. *usability engineering* ergäben sich dabei u.a. folgende Forschungsfragen (S. 3):

- „Wie benutzerfreundlich ist die vorliegende Software für die unterschiedlichen Zielgruppen hinsichtlich Effektivität, Effizienz, Zufriedenheit?“
- „Wie hoch ist die Akzeptanz des Basisdokuments für die unterschiedlichen Zielgruppen?“
- „Wird die Akzeptanz der verschiedenen Zielgruppen durch die Erfahrung mit dem vorliegenden elektronischen Basisdokument gesteigert?“

Als Schwierigkeit bei der Auswahl geeigneter Methoden zur Beantwortung dieser Fragen wird die schlechte zeitliche Planbarkeit des Projektverlaufs genannt: die

8. Evaluation im Reallabor Strukturvorgaben: Planung der Projektgruppe

Beteiligung von Richter:innen und Anwält:innen ist freiwillig und das Basisdokument wird ergänzend zum Normalbetrieb im Zivilprozess getestet (S. 4). Auch einige Erhebungsmethoden müssen für das Reallabor Strukturvorgaben ausgeschlossen werden: Feldbeobachtung und Beobachtung in kontrolliertem Umfeld etwa fallen durch mangelnde Ressourcen und Datenschutzprobleme aus dem möglichen Methodenkanon (S. 5). Für geeignet dagegen halten die Projektmitglieder semistrukturierte Interviews, Fokusgruppen und Fragebögen (S. 6 ff.). In einem *mixed methods approach* könnten dabei quantitative und qualitative Methoden und Auswertungen kombiniert werden (S. 7).

Im Hinblick auf das regulatorische Erkenntnisinteresse muss klar sein, „dass es bei Reallaboren um mehr geht als um die bloße Ermittlung der Gebrauchstauglichkeit bestimmter Anwendungsprogramme“ (S. 8). Ziel ist dabei die Formulierung von konkreten Handlungsempfehlungen für den Gesetzgeber – die sich nicht unbedingt nur konkret auf diese Implementierung des Basisdokuments beschränken müssen, sondern eben auch eine grundsätzliche Betrachtung des strukturierten Parteivortrags als abstrakter Idee vornehmen (S. 8). Auch in Bezug auf bisher vorgebrachte Bedenken soll die Evaluation mit tatsächlichen Nutzer:innen Antworten bringen (S. 9). Der Fokus der Evaluation soll „jedoch nicht auf einer detailreichen Austarierung und Weiterentwicklung der Software des genutzten Prototyps liegen – es handelt sich insofern nicht um eine typische Software-Evaluation –, sondern auf der Erlangung regulatorischer Aufschlüsse für die Zukunft“ (S. 9).

Zu Beginn der Erprobungsphase wurden vom Projektteam Umfragen erstellt, die bei Meldung eines Verfahrens an die beteiligten Anwält:innen und Richter:innen versandt werden. Die Umfrage enthält standardisierte Fragebögen zu technischem Interesse (*affinity for technology interaction*, ATI) und Technologieakzeptanz (*technology acceptance model*, TAM) und Fragen zum Verfahren, zur Arbeitsweise, zur Einstellung gegenüber dem Vorschlag eines Basisdokuments, zur Motivation für die Teilnahme, zu Erwartungen und Befürchtungen hinsichtlich der Erprobung und zu Wünschen für zusätzliche Prototypfunktionen. Außerdem will die Projektgruppe

8. Evaluation im Reallabor Strukturvorgaben: Planung der Projektgruppe

Interviews mit teilnehmenden Richter:innen und Anwält:innen führen.

Zu den beiden Forschungsfragen zur Akzeptanz werden mit den Umfragen Daten erhoben (vor allem quantitativ durch den TAM). Zur Bewertung der Nutzungsfreundlichkeit können die Interviews qualitative Eindrücke liefern.

Bis Dezember 2023 wurden der Projektgruppe etwa 50 Verfahren gemeldet, die Zahl der *Matches* liegt im niedrigen einstelligen Bereich. Interviews wurden bisher mit Richter:innen geführt, die aus der Erprobung ausgeschieden sind (z.B. weil sie von einem Testgericht an ein anderes gewechselt sind), und einem erprobenden Anwalt.

Diese Planungen bzw. Veröffentlichungen der Projektgruppe und ihre Überlegungen zu den Rahmenbedingungen des Reallabors sollen hier erwähnt werden, die Bewertung der Eignung der HCI-Evaluationsmethoden nehme ich im Folgenden aber grundsätzlich unabhängig von dieser Einschätzung vor.

9. Anwendung von Evaluationsmethoden auf das Reallabor Strukturvorgaben

Im Folgenden schildere ich mein Vorgehen bei der Analyse der Anwendung von Evaluationsmethoden auf das Reallabor Strukturvorgaben und die Ergebnisse dieser Analyse. Im Rahmen der Analyse wurde zur Bewertung der Methoden ein Kriterien- und Fragenkatalog entwickelt, der in Kapitel 9.1 vorgestellt wird. Die Ergebnisse der Analyse sind in Kapitel 9.2 dargestellt und in Kapitel 9.3 ausführlich diskutiert. Auf Basis dieser Diskussion gebe ich abschließend eine Empfehlung für eine Methodenkombination zur Evaluation im Rahmen des Reallabors Strukturvorgaben.

Zunächst sollen noch einige Begrifflichkeiten eingeordnet werden: Neben dem Begriff *Usability* ist inzwischen auch der Begriff der *User Experience* (UX) in der HCI weit verbreitet. Die beiden Konzepte sind nicht ganz trennscharf und werden in Studien auch häufig nicht explizit getrennt betrachtet. Allgemein lässt sich jedoch sagen, dass *Usability* ein Qualitätsmerkmal eines UI ist, während UX einen breiteren Ansatz abdeckt (Chesnut & Nichols, 2014; Norman & Nielsen, o.J.). Es geht bei UX nicht nur darum, wie einfach oder intuitiv ein UI zu nutzen ist, sondern um umfassendere Ziele wie Nützlichkeit und allgemeines Engagement. Zum Teil wird UX auch als Designpraxis verstanden (Chesnut & Nichols, 2014, S. 8), die sich auf die Schaffung von Erfahrungen konzentriert, die einfach zu benutzen und für die Nutzer:innen zufriedenstellend sind (Vermeeren et al., 2010, S. 521) – *Usability* stellt dabei eine Komponente von UX dar.

In der HCI werden sowohl für die *Usability* als auch für die *User Experience* Evaluationen durchgeführt. Für die folgende Betrachtung von Evaluationsmethoden werden deshalb beide Bereiche herangezogen. Dieser ganzheitlichere Ansatz scheint auch mit Blick auf die Projektziele und Erwartungen an die Evaluation im Reallabor sinnvoll (siehe Kapitel 5.2).

9. Anwendung von Evaluationsmethoden auf das Reallabor Strukturvorgaben

Zur Bewertung der Eignung einzelner Methoden für die Evaluation im Reallabor Strukturvorgaben habe ich eine Sammlung von Evaluationsmethoden erstellt und Aspekte bzw. Kriterien gesammelt, die für die Bewertung der Methode in Bezug auf deren Eignung für die Evaluation im Reallabor Strukturvorgaben wichtig sind. Diese Kriterien betreffen die Rahmenbedingungen im Reallabor Strukturvorgaben, menschliche und organisatorische Aspekte, UX-Metriken, die Projektziele und die Nutzung des Basisdokuments in zivilrechtlichen Verfahren. Die Kriterien sind in Kapitel 9.1 genauer beschrieben.

In Methodensteckbriefen habe ich für jede Methode beurteilt, wie gut sie für die Bewertung der einzelnen Aspekte geeignet ist. Die Methodensteckbriefe sind im digitalen Anhang hinterlegt (Ordner /4_Methodenanalyse) und online abrufbar⁶⁹, beispielhaft sind zwei Methodensteckbriefe (Heuristische Bewertung und Interview) im Anhang in Kapitel A.4 abgedruckt. Zur besseren Übersicht sind die Ergebnisse der Bewertung in einer Ergebnismatrix dargestellt (siehe Abbildung 6 in Kapitel 9.2).

Die Sammlung von Evaluationsmethoden basiert auf verschiedenen Quellen aus Literatursuche (z.B. Nielsen & Mack (1994); Vermeeren et al. (2010); Maia & Furtado (2016); Lazar et al. (2017); Shneiderman et al. (2017); Pettersson et al. (2018); Inan Nur et al. (2021)) und Internetrecherche⁷⁰. Methoden, die für die Evaluation einer Web- oder Desktop-Anwendung nicht geeignet sind, habe ich nicht in die Liste aufgenommen. Ebenso habe ich Methoden ausgeschlossen, die nicht für die Evaluation eines quasi voll-funktionalen Prototyps geeignet sind (z.B. *Technology Probes*⁷¹).

Die Bindung an die ERVV (Elektronischer-Rechtsverkehr-Verordnung) bildet grundsätzlich die einzige Einschränkung für den Basisdokumentprototyp – bei der Datei-

⁶⁹Link zu den Methodensteckbriefen: https://drive.google.com/file/d/1GIGNktE_OOJV8rGsg6a6j-0nrb3POGGi/view?usp=sharing

⁷⁰UX-Methodensammlungen finden sich z.B. auf <https://experienceresearchsociety.org/ux/evaluation-methods/> oder <https://www.uxmethods.guru/catalog>, aufgerufen am 16.11.2023

⁷¹Technology Probes: <https://www.interaction-design.org/literature/topics/technology-probes>, abgerufen am 16.11.2023

9. Anwendung von Evaluationsmethoden auf das Reallabor Strukturvorgaben

übermittlung ist für die Erprobung nicht die bestvorstellbare Lösung umsetzbar. Dieser Tatsache sollte in der Evaluation grundsätzlich Rechnung getragen werden können.

Im Folgenden wird zunächst der Kriterien- und Fragenkatalog vorgestellt, der die Rahmenbedingungen für die Evaluation im Reallabor Strukturvorgaben bildet. Anschließend sind die Ergebnisse der Analyse dargestellt und werden diskutiert.

9.1. Kriterien- und Fragenkatalog

Ob eine Methode für den Einsatz in der Evaluation des vorliegenden Reallabors Strukturvorgaben geeignet ist, möchte ich anhand verschiedener Kriterien bewerten. Nicht alle dieser Kriterien sind spezifisch für das vorliegende Reallabor, manche lassen sich insgesamt für das Reallabor als Methode anwenden (z.B. der Einsatz unter realen Bedingungen). Für diese Sammlung kann kein Anspruch auf Vollständigkeit erhoben werden, die zentralen Aspekte sind aber festgehalten. Die gesammelten Kriterien lassen sich verschiedenen Kategorien zuordnen, die im Folgenden beschrieben werden.

Rahmenbedingungen im Reallabor wie in Kapitel 3.4 und Kapitel 5.2 beschrieben.

- langer Erprobungszeitraum: Die Erprobungsphase ist auf 18 Monate angelegt – ist die Methode geeignet diesen Zeitraum zu evaluieren?
- lange Verfahren: Zivilrechtliche Verfahren können mehrere Jahre dauern – ist die Methode geeignet diesen langen Zeitraum zu evaluieren bzw. kann sie beliebig während des Verfahrens genutzt werden?
- unregelmäßige Nutzung des Prototyps: Im Verfahren nutzen Anwält:innen und Richter:innen den Prototyp nach Bedarf – leiden die Methode bzw. die Ergebnisse darunter, wenn Teilnehmer:innen den Prototyp nicht regelmäßig nutzen?
- verschiedene Nutzendengruppen: Anwält:innen und Richter:innen, sowie weitere Angestellte arbeiten mit dem Basisdokument (zum Teil im Prototyp, zum Teil mit den Basisdokumentdateien in anderer Justiz-Software) – kann die Methode die verschiedenen Gruppen einbeziehen?
- Freiwilligkeit der Teilnahme (Motivation): Teilnehmende Anwält:innen und Richter:innen sind vermutlich am Projekt interessiert und motiviert, es sind

9. Anwendung von Evaluationsmethoden auf das Reallabor Strukturvorgaben

bisher aber nicht sehr viele – kann die Methode auch mit wenigen Teilnehmer:innen genutzt werden?

- sensible Daten / Datenschutz: Das Basisdokument enthält persönliche Daten der Parteien, die dem Datenschutz unterliegen – ist der einzuhaltende Datenschutz ein Problem für die Nutzung der Methode?
- iterative Entwicklungen am Prototyp: Auch während der Erprobung und Evaluation wird der Prototyp weiterentwickelt – kann die Methode mit diesen Veränderungen am Prototyp umgehen?
- Einsatz im Realbetrieb: Das Basisdokument wird im Realbetrieb erprobt – kann die Methode den Einsatz im Realbetrieb evaluieren?

Menschliche und organisatorische Aspekte, ausgewählt nach Rupp et al. (2009, S. 97).

- geringe Motivation der Teilnehmer:innen: Leiden die Methode bzw. die Ergebnisse darunter, wenn Teilnehmer:innen wenig motiviert sind?
- schlechte kommunikative Fähigkeiten der Teilnehmer:innen: Leiden die Methode bzw. die Ergebnisse darunter, wenn Teilnehmer:innen ihre Probleme, Wünsche, Bedenken, etc. nicht gut ausdrücken können?
- implizites Wissen der Teilnehmer:innen: Ist die Methode geeignet, auch implizites Wissen („können, ohne sagen zu können, wie“) der Teilnehmer:innen zu erfassen?
- geringes Abstraktionsvermögen der Teilnehmer:innen: Verstehen die Teilnehmer:innen, dass die Anwendung ein Prototyp ist und können sie Probleme, Wünsche, Bedenken, etc. sinnvoll einordnen? Können Teilnehmer:innen vom Prototyp auf die Idee des Basisdokuments abstrahieren?
- divergierende Stakeholder-Meinungen: Kann die Methode die unterschiedlichen Meinungen von Richter:innen, Anwälte:innen, deren Angestellten und Mandant:innen abbilden?
- problematische Gruppendynamik: Leiden die Methode bzw. die Ergebnisse unter einer womöglich problematischen Gruppendynamik (Machtgefälle zwischen Angestellten, Gegenüberstehen in Streitfall, ...)?
- hohe Verteilung der Stakeholder: Ist die Methode geeignet, Ergebnisse von räumlich verteilten Teilnehmer:innen zu erheben?
- hohe Komplexität der Systemabläufe: Kann die Methode die komplexen Abläufe in Gerichtsverfahren angemessen berücksichtigen/betrachten/evaluieren?
- Erfahrung von Analytiker:innen mit der entsprechenden Ermittlungstechnik: Kann die Methode sauber angewandt werden? (immer subjektiv zu bewerten)

9. Anwendung von Evaluationsmethoden auf das Reallabor Strukturvorgaben

Diese Aspekte sind nicht spezifisch für das vorliegende Reallabor, sie lassen sich auf andere Projekte und Reallabore übertragen.

UX-Metriken, die in der Evaluation gängig sind (siehe z.B. Chesnut & Nichols (2014); Shneiderman et al. (2017) u.a.).

- Aufgabenperformanz der Teilnehmer:innen
- Zeitperformanz der Teilnehmer:innen
- Zufriedenheit der Teilnehmer:innen
- Konsistenz der Anwendung
- Ästhetik der Anwendung
- Aufgabenlast
- Technologieakzeptanz
- Erwartungen der Teilnehmer:innen
- Navigation in der Anwendung
- Benennungen / Nutzung von Fachbegriffen in der Anwendung
- Einfachheit der Nutzung
- Sichtbarkeit von Kernfunktionen
- Technologieperformanz
- Klarheit der Inhalte
- Identifikation der Fehler der Teilnehmer:innen
- Behandlung der Fehler der Teilnehmer:innen
- Zeit zum Erlernen der Handhabung
- Beibehalten des Gelernten im Laufe der Zeit

Auch diese Aspekte sind nicht spezifisch für das vorliegende Reallabor, sie lassen sich auf andere Projekte und Reallabore übertragen.

Neben diesen Kriterien, die die Methodenwahl beeinflussen, kann eine Methode auch besser oder schlechter geeignet sein, die Fragen zu beantworten, die für die Evaluation wichtig sind. Ich habe also auch einen **Fragenkatalog** erstellt, der im Rahmen der Evaluation genutzt werden kann bzw. soll. Quellen für diese Fragen sind meine Fokusgruppe zu Projektstand, Projektzielen und Erwartungen an die Evaluation und ein Planungsdokument der Projektgruppe für das Evaluationskonzept. Einzelne Fragen habe ich auf (quantifizierbare) Kriterien heruntergebrochen, für andere scheint das nicht sinnvoll möglich. Die Fragen betreffen verschiedene

9. Anwendung von Evaluationsmethoden auf das Reallabor Strukturvorgaben

Themenblöcke (Projektziel, Nutzung des Basisdokuments, Umsetzung des Basisdokuments, Bedenken bei der Nutzung des Basisdokuments, Arbeitsprozesse der Anwält:innen und Richter:innen) und sind spezifisch für das vorliegende Reallabor, womöglich sind einzelne Fragen aber auch für andere Projekte spannend. Der Fragenkatalog ist im Anhang hinterlegt (Kapitel A.2). Für die 27 einzelnen Methoden habe ich jeweils bewertet, ob sie insgesamt zur Beantwortung dieser Fragen geeignet sind.

9.2. Ergebnisse der Analyse

In Abbildung 6 sind die Ergebnisse der Analyse von 27 Methoden in Form einer Matrix dargestellt.⁷² Die Eignung einer Methode für ein Kriterium ist jeweils auf einer fünfstufigen Skala von -- bis ++ bewertet, die Felder der Tabelle sind entsprechend eingefärbt. Die Begründung für die Bewertung ist den 27 Methodensteckbriefen im digitalen Anhang (Ordner /4_Methodenanalyse) zu entnehmen (auch online hinterlegt⁷³). Die Matrix dient der besseren Übersichtlichkeit, exemplarisch sind aber auch zwei Steckbriefe im Anhang in Kapitel A.4 abgedruckt.

Mittelt man die Bewertungen pro Methode, ergibt sich ein Indikator für die Eignung der Anwendung im Reallabor Strukturvorgaben. Insgesamt positiv bewertet sind neun Methoden: *Contextual Inquiry* (CI, +), *Umfrage* (+), *Experience Sampling Method* (ESM, +), *UX Curve* (++) , *Tagebuchstudie* (++) , *Private Camera Conversation* (PCC, ++), *Fokusgruppe* (+), *Interview* (++) und *User Data Logging* (UDL, ++). Diese Methoden sind in Abbildung 6 hervorgehoben (fett).

Die Methoden, die in Kapitel 6 nicht vorgestellt sind, werden im Anhang (Kapitel A.3) beschrieben (CI, ESM, UX Curve, PCC, UDL).

⁷²Link zur Methoden-Matrix: <https://drive.google.com/file/d/10101Xr70yQyMa3v49PDo6jtOclSKmJLg/view?usp=sharing>

⁷³Link zu den Methodensteckbriefen: https://drive.google.com/file/d/1GIGNktE_OOJV8rGsgea6j-0nrb3POGGi/view?usp=sharing

9. Anwendung von Evaluationsmethoden auf das Reallabor Strukturvorgaben

		Kontrolliertes Experiment	Kontrollierte Beobachtung	Thinking Aloud	Feldbeobachtung/Shadowing	Co-Discovery	Contextual Inquiry	Umfrage (Fragebogen)	Experience Sampling Method	UX-Curve	Product Semantic Analysis	Tagebuchstudie	Private camera conversation	Fallstudie	Fokusgruppe	Interview	Expert Interview	Contextual Interview	UX laddering	Heuristische Bewertung	Konsistenzprüfung	Cognitive Walkthrough	Pluralistic Walkthrough	Guidelines Review	Formal Usability Inspection	User Journey	Automatisierte Tests	User Data Logging
Rahmenbedingungen Reallabor	Erprobungszeitraum	o	o	+	+	+	++	++	++	++	o	++	++	++	++	o	o	++	++	o	o	o	o	o	o	o	o	++
	Verfahren	o	o	+	o	+	++	++	++	++	o	++	++	-	++	++	o	++	++	o	o	o	o	o	o	o	o	++
	Prototypnutzung	o	o	--	--	--	--	--	--	--	o	++	++	--	++	++	o	o	o	o	o	o	o	o	o	o	o	++
	Nutzendengruppen	++	++	++	++	-	++	++	++	++	+	++	++	++	++	++	o	o	++	++	o	o	+	o	o	o	o	-
	Freiwilligkeit	--	--	--	+	--	-	-	-	-	-	--	--	--	--	--	o	o	--	--	o	o	o	o	o	o	o	++
	Datenschutz	++	++	--	--	--	+	++	++	++	++	++	++	--	++	++	++	o	++	++	o	o	++	++	++	++	++	+
	Entwicklungen	+	o	+	o	--	+	+	++	++	++	++	++	++	+	+	+	--	+	+	--	--	--	--	--	--	--	++
Realbetrieb	--	--	++	++	+	++	++	++	++	++	--	++	++	++	++	++	++	++	++	--	--	--	--	--	--	--	++	
Menschliche / organisatorische Aspekte	Motivation	-	-	-	++	+	+	--	--	-	--	--	-	+	+	+	++	+	o	o	o	o	-	o	o	o	o	++
	Kommunikation	o	++	-	++	-	++	++	o	o	--	--	--	++	++	++	++	++	++	o	o	o	o	-	o	o	o	o
	implizites Wissen	o	+	+	++	+	++	--	--	--	--	--	--	++	-	--	--	--	o	o	o	o	+	o	o	o	o	o
	Abstraktion	o	o	++	o	-	+	-	++	+	o	-	--	+	+	+	+	+	o	o	o	o	+	o	o	o	o	o
	Meinungen	o	++	++	o	--	++	++	o	++	++	++	++	++	++	++	++	++	++	++	o	o	++	o	o	o	o	o
	Gruppendynamik	++	++	++	++	--	++	++	++	++	++	++	++	++	--	++	++	++	++	++	o	o	o	--	o	o	o	o
	Verteilung	-	-	++	++	--	++	++	++	++	++	++	++	-	--	++	+	+	o	o	o	o	--	o	o	o	o	++
	Komplexität	--	--	+	++	++	-	o	--	--	--	--	--	++	++	++	++	+	-	o	o	o	++	o	o	o	o	o
	Erfahrung	++	--	+	+	--	+	+	+	--	--	--	--	--	--	--	--	--	--	--	--	--	+	--	+	--	--	+
	Aufgabenperformanz	++	+	--	+	--	o	--	o	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	+
UX-Metriken	Zeitperformanz	++	+	--	+	--	o	--	o	--	--	--	--	--	--	--	--	--	--	--	--	+	--	--	--	--	+	
	Zufriedenheit	++	o	+	o	+	+	+	-	+	+	+	+	++	++	++	++	++	-	+	++	++	++	+	+	+	--	
	Konsistenz	++	o	++	o	+	+	+	--	o	o	+	+	++	++	++	++	++	+	+	++	++	++	+	+	+	--	
	Ästhetik	++	o	++	o	+	+	+	--	o	o	+	+	++	++	++	++	++	+	+	--	--	--	--	+	+	+	--
	Aufgabenlast	++	o	+	o	+	o	+	o	o	+	+	+	+	+	+	+	+	--	--	+	+	+	+	+	+	--	
	Technologieakzeptanz	++	o	++	o	+	o	+	-	o	o	+	+	++	++	++	++	++	++	--	--	--	--	-	--	+	--	
	Erwartungen	++	o	+	o	+	+	++	--	o	--	+	+	++	++	++	++	++	++	--	--	--	--	-	--	--	--	
	Navigation	++	+	++	+	+	+	+	--	o	o	+	+	++	++	++	++	++	+	+	--	--	++	++	+	+	+	
	Benennungen	++	o	++	o	+	++	++	--	o	o	+	+	++	++	++	++	++	++	--	--	--	--	o	o	+	+	
	Einfachheit	++	o	++	o	+	++	+	-	+	o	+	+	++	++	++	++	++	++	--	--	+	+	+	+	+	+	
	Kernfunktionen	++	+	++	+	+	++	++	--	o	o	+	+	++	++	++	++	++	++	+	+	--	++	++	+	+	+	
	Technologieperformanz	++	+	++	+	+	++	--	o	o	+	+	+	++	++	++	++	++	++	+	+	--	+	-	+	+	+	
	Klarheit	++	o	++	o	+	++	--	o	o	+	+	+	++	++	++	++	++	++	+	+	--	+	++	+	+	--	
	Fehleridentifikation	++	+	++	+	+	++	++	--	o	--	+	+	+	+	+	+	+	+	--	--	+	-	-	+	+	+	
	Fehlerbehandlung	++	+	++	+	+	++	++	--	o	--	+	+	+	+	+	+	+	+	--	--	+	-	-	+	+	+	
Lernen	++	o	+	o	--	+	+	--	o	o	+	+	+	+	+	+	+	+	--	--	+	--	-	+	--	--		
Beibehalten	-	o	+	o	+	+	+	--	o	o	+	+	+	+	+	+	+	+	--	--	--	--	-	--	--	--		
Evaluationsfragen	-	--	-	--	-	++	++	--	--	+	+	+	++	++	++	++	++	++	o	o	o	o	o	o	o	o	o	

Abbildung 6: Ergebnisse der Analyse der Eignung von UX-Methoden für bestimmte Kriterien (hier abgekürzt). Die Eignung einer Methode für ein Kriterium ist auf einer fünfstufigen Skala (–– bis ++) bewertet, die Felder sind entsprechend eingefärbt. Methoden, deren Bewertung insgesamt positiv ist, sind hervorgehoben (fett). Eigene Abbildung.

9.3. Diskussion der Analyseergebnisse

Im Folgenden wird die Anwendung der neun positiv bewerteten Methoden im Reallabor Strukturvorgaben diskutiert und eine Empfehlung für eine Methodenkombination zur Evaluation gegeben. Als Beispiel einer negativen Bewertung möchte ich außerdem darstellen, warum die heuristische Bewertung nach meiner Analyse

9. Anwendung von Evaluationsmethoden auf das Reallabor Strukturvorgaben

so schlecht für das Reallabor Strukturvorhaben geeignet ist. Zum einen liegt das daran, dass die heuristische Bewertung losgelöst vom Reallabor (also eben nicht im echten Nutzungskontext) stattfindet, zum anderen daran, dass die gängigen Heuristiken (z.B. Nielsen (1994)) häufig nur die *Usability* eines Systems bewerten. So kann mit dieser Methode die Situation im realen Kontext nicht evaluiert werden, dadurch haben die Rahmenbedingungen des Reallabors teils auch keinen Einfluss auf die Anwendung der Methode. Es werden keine echten Nutzer:innen verschiedener Gruppen einbezogen, wodurch menschliche oder organisatorische Aspekte ebenfalls keinen Einfluss auf die Anwendung der Methode haben (weder negativ noch positiv). Weil der Fokus häufig auf der *Usability* liegt, werden andere UX-Metriken von den meisten Heuristiken nicht bewertet. So können auch nur die Evaluationsfragen beantwortet werden, die die *Usability* betreffen. Die meisten Fragen zur Nutzung des Basisdokuments können von Expert:innen in einer heuristischen Evaluation jedoch nicht bewertet werden. Insgesamt ist die heuristische Evaluation für den Einsatz im Rahmen der Evaluation eines Reallabors also wenig geeignet. Der Methodensteckbrief zur heuristischen Evaluation ist in Kapitel A.4 angehängt.

Nach dieser Einordnung eines Negativbeispiels möchte ich nun näher auf die neun positiv bewerteten Methoden eingehen und ihre Eignung jeweils detaillierter kommentieren.

Tagebuchstudie und PCC werden ohne Eingriffe einer:ines Forschenden von den Teilnehmer:innen durchgeführt. Beide Methoden erfordern regelmäßige Einträge bzw. Aufnahmen und sind für die Teilnehmer:innen grundsätzlich – aber eben auch nach eigenem Ermessen – zeitintensiv. Sie bieten das Potenzial, Daten über einen längeren Zeitraum zu sammeln und einen Eindruck davon zu bekommen, wie die Anwendung im realen Kontext tatsächlich genutzt wird. Die Qualität und Quantität der Ergebnisse hängt stark von der Motivation der Teilnehmer:innen ab und lässt sich von den Forschenden nicht kontrollieren. Im Reallabor Strukturvorgaben scheint es sehr unwahrscheinlich, dass Teilnehmer:innen für diese Art der Evaluation gewonnen werden können – schon ohne weiteren Arbeitsaufwand durch eine Evaluation ist die Zahl der Testwilligen gering.

9. Anwendung von Evaluationsmethoden auf das Reallabor Strukturvorgaben

Auch das Ausfüllen von Umfragen (in Form von selbsterstellten Fragebögen, ESM oder *UX Curve*) erfolgt unmoderiert. Es lassen sich dafür aber womöglich leichter bzw. mehr Teilnehmer:innen gewinnen, da die Umfrage einfach versandt werden kann. Für ESM müssen Teilnehmer:innen allerdings nach jeder Nutzung des Basisdokuments einen kurzen Fragebogen ausfüllen. Es scheint fraglich, ob die Rücklaufquote für Umfragen besonders hoch wäre. Auch wie ordentlich bzw. ausführlich die Umfrage ausgefüllt wird, hängt stark von der Motivation der Teilnehmer:innen ab. Die Quantität und Qualität der Ergebnisse kann also auch mit dieser Methode stark variieren.

Bei CI, Fokusgruppe und Interview treten Forschende direkt in Kontakt mit Teilnehmer:innen und können bei Unklarheiten oder Problemen nachfragen. Die Sitzungen können mehr oder weniger stark strukturiert bzw. moderiert werden, zentrale Themen können auf jeden Fall angesprochen, kommentiert oder diskutiert werden. Eine Fokusgruppe scheint im Reallabor Strukturvorgaben schwierig zu realisieren, da alle Gruppen von Nutzer:innen beteiligt werden sollten – neben zeitlichem bedeutet das auch organisatorischen Aufwand für die Teilnehmer:innen: Anwält:innen und Richter:innen, im Optimalfall auch weitere Angestellte des Gerichts oder der Kanzlei und eventuell Mandant:innen, müssen zu einer (mehrstündigen) Sitzung versammelt werden. Eine problematische Gruppendynamik kann dazu führen, dass die Ergebnisse knapp ausfallen oder einzelne Gruppen von Nutzer:innen nicht adäquat abgebildet sind. Um gezielt Rückmeldung zu einzelnen Features oder verschiedenen Designvorschlägen zu bekommen, scheint eine Fokusgruppe geeignet – für die ausführliche Evaluation des Basisdokuments im Reallabor Strukturvorgaben aber eher nicht. Allein der Umfang des Fragenkatalogs lässt sich in einer Gruppendiskussion kaum abdecken und müsste stark reduziert werden.

Feedback von einzelnen Personen zu ihren Erfahrungen mit dem Basisdokument zu sammeln, scheint mir vielversprechender. Die Wahl zwischen *Contextual Inquiry* und leitfadengestütztem Interview fällt aber weniger eindeutig aus. Gerade im Reallabor spielt der Kontext der Nutzung eine zentrale Rolle. Problematisch können vor allem Datenschutzfragen sein und ob sich Teilnehmer:innen dabei wohlfühlen,

9. Anwendung von Evaluationsmethoden auf das Reallabor Strukturvorgaben

eine:n fremde:n Forschende:n in ihr Arbeitsumfeld zu lassen und die eigene Arbeit zu zeigen. Die CI muss außerdem erfolgen, während der Prototyp genutzt wird, was Schwierigkeiten für die Organisation der Sitzung bedeutet. Ich halte Interviews nach meiner Analyse für sehr gut geeignet, um das Reallabor Strukturvorgaben bzw. den Einsatz des Basisdokuments zu evaluieren und möchte für eine flexible Lösung plädieren: wenn die Teilnehmer:innen einverstanden sind, sollte das Interview am Arbeitsplatz stattfinden. So können Probleme, Arbeitsabläufe, etc. direkt anhand des Basisdokuments (Prototyp oder Dateien) von Teilnehmer:innen illustriert und kommentiert werden. Gerade für Aspekte der *Usability* scheint mir das wichtig und sinnvoll zu sein. Gleichzeitig sollen die Fragen der Evaluation (siehe Fragenkatalog im Anhang, Kapitel A.2) gestellt werden können, sodass der Fokus nicht nur auf der Beobachtung der Nutzung des Basisdokuments liegt.⁷⁴

Wie *User Data Logging* im Reallabor Strukturvorgaben eingesetzt werden könnte, möchte ich an dieser Stelle ebenfalls etwas ausführlicher darstellen und diskutieren, weil der Modus der Datengewinnung sich sehr von den anderen vorgestellten Methoden unterscheidet. Im Basisdokumentprototyp könnte es mit Blick auf die Evaluationsfragen sinnvoll sein, folgende Daten zu loggen: (a) Zeitpunkt der Prototypnutzung (um Informationen zur Verfahrensdauer bzw. der Dauer bis zur mündlichen Verhandlung zu gewinnen), (b) bei Erstellung einer neuen Version des Basisdokuments (als Indikator für die Anzahl der Satzwechsel) und (c) Länge der Sätze/Beiträge der Parteien oder des Basisdokuments zu diesem Zeitpunkt (als Indikator dafür, ob die Strukturierung auch für kürzere Sätze sorgt), (d) wenn die angebotenen Strukturierungsmöglichkeiten (v.a. Bezugnahme, Anlegung einer eigenen Sortierung) genutzt werden (als Indikator dafür, ob das Basisdokument so eingesetzt wird wie geplant), (e) wenn Vortrag als streitig oder unstreitig markiert wird (als Indikator dafür, ob diese Funktion genutzt wird, ob es weniger Streitpunkte gibt und ob Richter:innen auf das Anlegen einer eigenen Relationstabelle verzichten) und (f) wenn (zu definierende) Kernfunktionen

⁷⁴Auch bei einem *Contextual Interview* liegt der Fokus auf der Beobachtung von Nutzer:innen und ihren Interaktionen, sodass diese Methode in meiner Bewertung schlechter abschneidet als ein leitfadengestütztes Interview.

9. Anwendung von Evaluationsmethoden auf das Reallabor Strukturvorgaben

(z.B. Suchfunktion, Hilfefunktion, ...) genutzt werden (als Indikator dafür, ob der angebotene Funktionsumfang genutzt wird bzw. sinnvoll ist). Diese Listung ist meine persönliche Einschätzung dazu, welche Daten für die Beantwortung der Evaluationsdaten hilfreich sein könnten. Welche Daten wirklich geloggt werden dürfen, müsste innerhalb der Projektgruppe (v.a. auch mit Beteiligung der juristischen Expert:innen) genauer geprüft und besprochen werden. In der aktuellen Version des Basisdokumentprototyps wird keinerlei Logging der Nutzungsdaten betrieben. Es besteht überhaupt kein Datentransfer vom Prototyp zu den Forschenden. Für Forschende ist grundsätzlich jede Interaktion der Nutzer:innen spannend, mir scheint ein umfassendes Logging durch den Einsatz im Reallabor (im Gegensatz zu einer Laborstudie) wegen der geschilderten Datenschutzproblematik und der Frage der Arbeitsüberwachung am Arbeitsplatz aber kaum vertretbar – auf jeden Fall müsste die explizite Zustimmung der Teilnehmer:innen eingeholt werden und die Teilnehmer:innen über das Logging aufgeklärt werden. Trotzdem können Interaktionsdaten sicher Einblicke dazu geben, wie der Basisdokumentprototyp genutzt wird. Mir scheint der Einsatz von UDL auch deswegen sinnvoll und wichtig, weil die Bewertung einer momentanen Nutzungserfahrung häufig nicht sehr zuverlässig ist, um die Erfahrung der Nutzer:innen im wirklichen Leben oder den Erfolg eines Produkts zu bewerten (Kujala et al., 2011; Maia & Furtado, 2016). Auch ein Interview ist eine Momentaufnahme und die Aussagen der Interviewten beruhen auf retrospektiven Bewertungen des Langzeitexperiments (Kujala et al., 2011). Geloggt werden Daten dagegen während des gesamten Erprobungszeitraums, sodass sie Aufschluss über die tatsächliche Nutzung geben.

Wie in jedem Projekt spielt auch im Reallabor Strukturvorgaben die Frage der Ressourcen eine Rolle bei der Planung der Evaluation. Deswegen möchte ich auch diesen Aspekt in meine Bewertung einbeziehen bzw. kommentieren.

Letztlich möchte ich auf Grundlage meiner Bewertung und Argumentation – wie in der HCI-Forschung üblich – im Sinne einer Triangulation (Pettersson et al., 2018) eine Kombination verschiedener Methoden vorschlagen: zum einen *User Data Log-*

9. Anwendung von Evaluationsmethoden auf das Reallabor Strukturvorgaben

ging einzusetzen, um über den gesamten Zeitraum Daten zur Interaktion mit dem Basisdokumentprototyp zu sammeln, zum anderen zu einem späten Zeitpunkt im Erprobungszeitraum Interviews mit beteiligten Nutzer:innen zu führen und sie zum Einstieg in das Interview ihre Erfahrungen mit dem Basisdokument in der *UX Curve* bewerten zu lassen. Sowohl die *UX Curve* als auch eine Auswertung der Logging-Daten (wenn der:die Befragte zur Gruppe der Prototypnutzenden gehört) kann im Interview als Grundlage für vertiefende Fragen und Auskünfte genutzt werden. In Ergänzung zu den Fragen des Fragenkatalogs kann meiner Meinung nach so ein umfängliches Bild der Erprobung des Basisdokumentprototyps gewonnen werden. Auch die Tatsache, dass der Dateiaustausch durch die Bindung an die ERVV die größte Abweichung des Prototyps von einer Ideallösung ist, kann durch Interviews hoffentlich adäquat abgebildet werden. Für die Projektgruppe ist klar, dass der Datenaustausch so nicht optimal ist und in einem finalen System anders umgesetzt werden müsste. In Interviews könnten die Teilnehmer:innen auch darauf Bezug nehmen und Rückmeldung zur zukünftigen Gestaltung geben.

Auch in Hinblick auf den Aspekt der im Projekt vorhandenen Ressourcen scheint mir diese Methodenkombination sinnvoll: Das Logging muss geplant und implementiert werden, läuft während der Erprobung dann aber im Hintergrund. Für die *UX Curve* sind Vorlagen vorhanden, der Einsatz bedeutet also kaum Zeitaufwand in der Planung. Planung, Durchführung und Auswertung der Interviews sind dagegen wesentlich zeitaufwendiger. Vor diesem Hintergrund sollte eine Reduzierung des Fragenkatalogs in der Projektgruppe diskutiert werden – so lässt sich die zeitliche Belastung sowohl für die Interviewten als auch die Interviewenden und die anschließende Auswertung reduzieren. Auch mit den Überlegungen der Projektgruppe Althammer et al. (2023) scheint mir die vorgeschlagene Methodik kompatibel.

10. Diskussion und Fazit

Im Folgenden sollen die Ergebnisse dieser Arbeit diskutiert und ein Fazit gezogen werden.

Dass ich nach Monaten der Erprobung eine Methodenkombination zur Evaluation vorschlage, die man vor Beginn der Erprobung hätte implementieren müssen, ist offensichtlich nicht optimal. Die Publikation der Projektgruppen mit Ideen für die Evaluationsmethodik erfolgte im März mit Beginn der Erprobung (Althammer et al., 2023), die Planung fand also schon vor Beginn der Erprobung statt. Allerdings gab es zu diesem Zeitpunkt wohl noch keine konkrete Implementierung der Evaluationsmethoden, die wurde dann projektbegleitend entwickelt – für Interviewleitfäden, die erst nach einer Zeit der Erprobung genutzt werden, funktioniert das gut. Für die Gewinnung von Interaktionsdaten innerhalb des Prototyps allerdings kaum, sodass die konkrete Planung der Evaluation in der Vorbereitungsphase dafür notwendig ist. Durch die iterative Weiterentwicklung am Prototyp hätten sich sogar später noch Zeitpunkte zur Implementierung ergeben – auch in diesem Sinne zeigt sich eine Weiterentwicklung während der Erprobung als sinnvoll.

Insgesamt scheint mir ein iteratives Vorgehen (wie es im Ansatz des *User Centered Design* beschrieben ist) nicht nur für die Systementwicklung, sondern auch für solche längerfristigen Projekte sinnvoll: zu Beginn des Projekts können Leitfäden wie der von Bergmann et al. (2005) als Stütze für die Planung der Projektarbeit dienen (um Ziele, Strukturen, Kommunikation, Aufgaben oder Rollen festzuschreiben), im Laufe des Projekts kann diese Planung dann reflektiert und eventuell angepasst werden. Die Zwischenevaluation in Form der Gruppendiskussion wurde von den Beteiligten sehr positiv kommentiert und als hilfreich wahrgenommen, um sich den aktuellen Stand mit seinen Problemen und die Ziele des Projekts erneut zu verdeutlichen. Letztlich verstehe ich diese Art von Reflexion als Teil des Projektma-

nagements – das natürlich Ressourcen benötigt, aber meiner Meinung nach Vorteile für alle Beteiligten bringt.

Gerade um grundlegende Probleme aufzudecken (hier z.B. die Bedenken der Anwält:innen, dass ihnen eine Struktur für den Parteivortrag vorgegeben wird) und ihnen sinnvoll zu begegnen, scheint mir ein solch reflektierendes Vorgehen sinnvoll. In der Literatur wird deutlich, dass die Idee des Basisdokuments sehr kontrovers diskutiert und kritisch gesehen wird – dass diese Kritik sich häufig aber nicht auf die konkrete Implementierung im Prototyp beziehen, sondern eben auf die abstraktere Idee bzw. die verschiedenen Vorschläge. Das deutet für mich darauf hin, dass es eine gute Entscheidung war, das *Wording* bzw. *Framing* des Reallabors auch im Projektverlauf noch zu ändern. Der Projekttitle „Strukturvorgaben für den Parteivortrag“ verstärkt das bekannte Vorurteil der Anwält:innen und scheint daher vorbelastet und eher unglücklich gewählt. Eine strukturierte Auseinandersetzung mit den Vorbehalten potenzieller Nutzer:innen und die Erarbeitung einer Kommunikationsstrategie helfen womöglich dabei, Projekte zu Beginn einer Erprobung positiver erscheinen lassen – denn eine Art Pilotprojekt, wie es ein Reallabor ist, ist wohl immer mit erhöhtem Arbeitsaufwand für die Beteiligten verbunden, die mit einer neuen Technologie konfrontiert sind und trifft deswegen häufig auf Ablehnung. Ob frühere Maßnahmen einen positiven Effekt auf die Teilnehmezahlen von Anwält:innen im Reallabor Strukturvorgaben gehabt hätten, lässt sich nicht abschätzen.

Das in der Fokusgruppe kommunizierte Ziel des Projekts, zu einem tatsächlichen Ergebnis in der Diskussion um die Strukturierung des Parteivortrags zu gelangen, ist im Rahmen der aktuellen Erprobung mangels Teilnehmer:innen womöglich nicht zu erreichen. Voraussichtlich bleibt die Stichprobe bis Ende der Projektlaufzeit sehr klein, sodass Rückmeldungen eher Einzelmeinungen bleiben und es schwierig ist, ein umfängliches Bild zu erhalten. Trotzdem können durch die Rückmeldungen (auf verschiedenen Wegen) sicher viele Einsichten für das Projekt und für die Implementierung des Basisdokuments gewonnen werden. Sowohl die Erfahrungen im Reallabor, als auch die Erkenntnisse zum Basisdokumentprototyp sind für die

Weiterführung des Projekts hilfreich.

Ich möchte auch kurz mein persönliches Fazit zur Arbeit im Projekt mit anbringen: Ich habe es sehr genossen, Einblicke in diese transdisziplinäre Forschung zu bekommen – beide Seiten haben viel voneinander und von den Praktiker:innen gelernt. Ich bin insgesamt sehr viel Offenheit begegnet, sowohl bei den Interviews mit Richter:innen im Praxisseminar 2021, als auch beim Besuch einer Geschäftsstelle 2023. Die Arbeit mit der Projektgruppe habe ich als stets konstruktiv und problemorientiert erlebt. Besonders positiv war für mich, wie begeistert die Jurist:innen von unserer Herangehensweise mit HCI-Methodik bzw. deren Ergebnissen waren und sind: wie wir strukturiert Anforderungen erheben und daraus Designvorschläge ableiten oder wie wir formale Richtlinien zur Reflexion nutzen.

Meine Arbeit soll zur Reflexion des Projekts beitragen und einen Leitfaden für die Analyse und Bewertung der Eignung von HCI-Methoden darstellen – auch über den Kontext des Reallabors Strukturvorgaben hinaus für andere transdisziplinäre Projekte. Der Kriterienkatalog kann für andere Projekte genutzt oder angepasst werden, viele der Aspekte treffen so oder ähnlich auch auf andere Projekte zu (z.B. lange Zeiträume, unregelmäßige Nutzung der neuen Technologie, verschiedene Nutzendengruppen, Datenschutz, der Einsatz im Realbetrieb oder menschliche Aspekte wie geringe Motivation, schlechte kommunikative Fähigkeiten und problematische Gruppendynamik). In anderen Projekten macht es womöglich Sinn, die Aspekte anders zu gewichten, zu ergänzen oder zu reduzieren. Meine Sammlung von Aspekten und Evaluationsmethoden kann jedoch als Ausgangspunkt dienen. Die Methodenmatrix (Abbildung 6) bietet dabei eine Übersicht der Bewertungen. So kann sie als Orientierungshilfe dienen und genutzt werden, um geeignete Methoden für die Evaluation von Reallaboren und anderen Projekten zu ermitteln.

Für den Abschlussbericht der Projektgruppe kann neben meinen Empfehlungen zur Evaluation auch meine Arbeit zur Erhebung des Projektstands und zentraler Probleme genutzt werden. Ich habe zu verschiedenen Themen Wissen gesammelt und kondensiert, – zum Projektverlauf, zur Methode Reallabor, zur Wahrnehmung des Projekts Basisdokument, zum Projektstand im Reallabor Strukturvorgaben, zu

Forschungsmethoden der Mensch-Maschine-Interaktion, zur Evaluation im Feld – um darauf meine Analyse der Evaluationsmethoden zu stützen. Auch dieses kondensierte Wissen kann für andere Projekte nützlich sein, besonders womöglich für andere Reallabore, an denen keine Forschenden der HCI beteiligt sind, die aber ebenfalls innovative Technologien erproben und evaluieren. Die Methodenmatrix soll dabei auch für Fachfremde eine Orientierungshilfe bieten.

Zuletzt möchte ich hier noch kurz die Rolle und Möglichkeiten von Prototypen reflektieren. Chesnut & Nichols (2014, S. 221) beschreiben Prototypen als Repräsentation der finalen Nutzungserfahrung. Im Reallabor Strukturvorgaben ist das für den Basisdokumentprototyp allerdings nur eingeschränkt der Fall. Durch die Bindung an den Datenaustausch im elektronischen Rechtsverkehr ist für diese Nutzung die Repräsentation des Prototyps nicht final – es ist allen Beteiligten klar, dass der Datenaustausch so nicht optimal ist und in einem finalen System anders umgesetzt werden müsste.

Einen anderen blinden Punkt der HCI-Evaluationsmethodik zeigen Salovaara et al. (2017) auf: Die Tatsache, dass ein Prototyp in einer Studie nie vollständig bewertet werden kann, wird selten berücksichtigt – der untersuchte Prototyp existiert in der gegenwärtigen Welt, unter den in der Studie geschaffenen Umständen, aber sein tatsächlicher Nutzungskontext ist ein teilweise unbekannter zukünftiger Zustand. Prototypen seien per definitionem Technologien, die dazu dienen, die Zukunft zu gestalten und zu erfahren, wie sie mithilfe der Technologie verändert werden könnte. Die empirische Bewertung füge jedoch noch etwas hinzu: „die Möglichkeit, etwas über die Fähigkeit zu lernen, die Welt zu verändern“ – bei der Evaluation geht es also auch darum, was aus dem Prototyp werden könnte (S. 2064). Im Reallabor bewegt man sich möglichst nahe an der realen Nutzung und kann „possible futures“ recht gut abbilden, aber eben nicht in allen Bereichen der Anwendung (bzw. des Basisdokumentprototyps).

Um die Frage der Zukunftsgestaltung geht es auch in der Gestaltung von Politik. Nach Sanchez & Epp (2023) „setzt [HCI] Zukunftsvisionen in der Gegenwart um“ (S. 1). Sie plädieren dafür, dass HCI und Politik voneinander lernen und profitieren

können: HCI könne eine partizipative, nachhaltige und wertorientierte Perspektive bieten, solle aber, um die Auswirkungen der Technologie in einer langfristigen und unsicheren Zukunft besser zu verstehen, neben der Technologie auch andere treibende Kräfte der Gesellschaft berücksichtigen. Die Politikgestaltung könnte von einem flexibleren und reaktionsfähigeren Evaluierungsumfeld profitieren und HCI-Studien könnten die Wirkung ihrer Ergebnisse über ihre eigene Disziplin hinaus verbessern.

Genau dies könnte auch die Chance für Reallabore sein, für die die Frage der Regulierung eine zentrale Rolle spielt.

11. Ausblick

Ein juristischer Ansatz, auf den ich während meiner Arbeit gestoßen bin, ist die Gesetzesfolgenabschätzung (GFA). Nach § 43 Absatz 1 Nr. 5 und § 44 Absatz 1 der Gemeinsamen Geschäftsordnung der Bundesministerien⁷⁵ (GGO) sind die Ressorts seit 2000 verpflichtet, die voraussichtlichen Gesetzesfolgen einer Regelung darzustellen. Analog dazu gibt es für Technologien die Technikfolgenabschätzung (TFA), die in den 1960/70er Jahren ebenfalls von der Politik ins Leben gerufen wurde (Grundwald, 2022, S. 17f). Methodisch führt beides zu weit, passt nicht exakt zur Aufgabenstellung und sprengt den Rahmen dieser Arbeit – aber dass es in den beiden Fächern, die gemeinsam am Projekt Basisdokument bzw. Strukturvorgaben arbeiten, Methoden zur Folgenabschätzung gibt, möchte ich hier zumindest ausblicksweise kommentieren. Die Gesetzesfolgenabschätzung umfasst „beabsichtigte Wirkungen und unbeabsichtigte Nebenwirkungen“, stellt dar, „ob die Wirkungen des Vorhabens einer nachhaltigen Entwicklung entsprechen“ und insbesondere auch, „welche langfristigen Wirkungen das Vorhaben hat“ (BMI, 2009, S. 3). Sie ist ein wichtiger Beitrag zur Verwaltungsmodernisierung und Rechtsoptimierung (trägt zur Verminderung der Regelungsmenge, zum sparsamen Umgang mit knappen Ressourcen und zur Vermeidung von Akzeptanzverlusten bei) (Böhret & Konzen-dorf, 2000, S. 6). Dazu scheint mir das Reallabor als Instrument besonders gut zu passen, da es eine Erprobung von Regulierung im realen Kontext vorsieht. Konkret kann das Reallabor Strukturvorgaben durch die Erprobung womöglich auch wichtige Einsichten für eine Gesetzesfolgenabschätzung eines digitalen Basisdokuments liefern. Analog zur GFA ist der Anspruch der TFA „Orientierungs- und Prognosewissen“ zu liefern, „indem sie Einschätzungen von Experten ebenso wie von betrof-

⁷⁵Gemeinsame Geschäftsordnung der Bundesministerien: Informationen und Download unter <https://www.bmi.bund.de/DE/themen/moderne-verwaltung/verwaltungsmodernisierung/geschaeftsordnung-bundesministerien/geschaeftsordnung-bundesministerien-node.html>, abgerufen am 19.12.2023

fenen Anspruchsgruppen, Trendprognosen, Machbarkeitsstudien und Wechselwirkungsanalysen einholt sowie konkrete Szenarien auf ihre Folgen hin durchspielt“ (Häußling, 2019, S. 361). Auch hier klingt für mich eine Nähe zum Reallabor an, sodass im weiteren Projektverlauf für das digitale Basisdokument neben der GFA eben auch die TFA spannende Einblicke liefern könnte.

Neben meiner hier vorliegenden haben sich auch andere Abschlussarbeiten am Lehrstuhl Medieninformatik mit dem Projekt Strukturvorgaben/Basisdokument beschäftigt.⁷⁶ Zum Beispiel wurden in einer Arbeit die Prozesse in Anwaltskanzleien erhoben und in die Entwicklung eines weiteren Basisdokumentprototyps einbezogen, in anderen Arbeiten wurde bzw. wird die Integration von Word, Latex und künstlicher Intelligenz in das Basisdokument untersucht bzw. ermöglicht. Für die Weiterentwicklung bzw. die Erkenntnisse zum Prototyp sind diese Arbeiten sicher hilfreich. Außerdem zeigen sie auf, wie komplex das Thema ist und dass es noch so viel mehr Möglichkeiten in der Umsetzung gibt. Das hat eventuell Auswirkungen auf die Evaluation, für die kann meine Arbeit bei der Neubewertung eine Orientierungshilfe bieten.

Wird für das Projekt tatsächlich eine Experimentierklausel beantragt und eine Erprobung kann im Rahmen der Justizplattform umgesetzt werden, hätte das nicht nur Auswirkungen auf den einzusetzenden Basisdokumentprototyp, sondern auch auf die Evaluation. Zum einen müsste ein Evaluationskonzept vorgelegt werden, zum anderen müssen veränderte Rahmenbedingungen berücksichtigt werden. Für beides kann meine Arbeit eine Basis bilden: Die Bewertung der Methoden und die Empfehlung der Methodenkombination (*User Data Logging*, *UX Curve* und Interviews) kann für die Erarbeitung des Evaluationskonzepts genutzt werden, die veränderten Rahmenbedingungen können analog zu meinen erarbeitet und bewertet werden. Für die Bewertung der Evaluationsmethoden bedeutet meiner Meinung nach eine Experimentierklausel womöglich aber gar keine große Veränderung der Rahmenbedingungen – denn auch wenn die Teilnahme an der Erprobung nicht frei-

⁷⁶Abschlussarbeiten am Lehrstuhl Medieninformatik im Rahmen des Projekts Strukturvorgaben/Basisdokument: https://wiki.mi.ur.de/arbeiten/inbearbeitung?dataflt%5B0%5D=stichworte_%3DBasisdokument, abgerufen am 19.12.2023

willig ist, wird es die Teilnahme an der Evaluation vermutlich sein. Trotzdem ist womöglich nicht zu unterschätzen, dass durch den Zwang zur Nutzung auch mehr Teilnehmer:innen ihre Meinung, Probleme oder Eindrücke kommunizieren möchten. Das würde für die Evaluation womöglich bedeuten, dass ein niederschwelliges Angebot zur Rückmeldung (vielleicht sogar innerhalb der Basisdokumentsoftware) sinnvoll zu integrieren wäre. Durch den Zwang zur Nutzung ist die Frage des Einsatzes von *User Data Logging* vermutlich noch kritischer zu beleuchten, sodass die Integration von Feedback direkt in der Software diese Methode ersetzen könnte. Der Basisdokumentprototyp müsste insgesamt an die neuen Bedingungen angepasst werden: der Versand der Dateien wäre wohl nicht mehr beschränkt durch die ERVV, stattdessen (im Optimalfall) die Integration der Basisdokumentsoftware in die Justizplattform umgesetzt – sodass gar keine Dateien mehr versendet werden müssten, sondern eine abwechselnd cloud-basierte Bearbeitung möglich ist. So ist der Ablauf insgesamt vereinfacht (bzw. konzentriert auf die Erstellung, Ergänzung, Strukturierung des Parteivortrags) und das Basisdokument ist besser in bekannte Systeme eingebunden.

Literaturverzeichnis

- Alavi, H. S., Lalanne, D. & Rogers, Y. (2020, April). The Five Strands of Living Lab: A Literature Study of the Evolution of Living Lab Concepts in HCI. *ACM Transactions on Computer-Human Interaction*, 27 (2), 1–26. Zugriff auf <https://www.doi.org/10.1145/3380958>
- Albrecht, J. (2021, Februar). „Modernisierung des Zivilprozesses“ - nicht an den Bürger:innen vorbei. *Anwaltsblatt. ZPO-Blog*. Zugriff am 2023-04-28 auf <https://anwaltsblatt.anwaltverein.de/de/zpoblog/zivilrichtertag-modernisierung-des-zivilprozesses-julian-albrecht>
- Althammer, C., Bauer, J., Böhm, V., Fehle, J., Mielke, B. & Wolff, C. (2023). Das Basisdokument geht ins Reallabor: zur Evaluation des Einsatzes bei Gericht. *Jusletter-IT* (März 2023). Zugriff auf <https://www.doi.org/10.38023/ba7351cd-d1dc-46d6-a62c-0b014b5f69e3>
- Anduschus, P.-O., Bienzeisler, B. & Prochazka, V. (2023). Innovationsmethode Reallabor: Eine Typologie. *Fraunhofer-Publica*. Zugriff auf <https://www.doi.org/10.24406/PUBLICA-1113> (Publisher: Fraunhofer IAO)
- Apraiz Iriarte, A. & Lasa Erle, G. (2020). Questionnaires for the evaluation of user experience and usability in industrial environments: a bibliographic review. *DYNA New Technologies*, 7 (1). Zugriff auf <https://www.doi.org/10.6036/NT9845>
- Arbeitsgruppe „Modernisierung des Zivilprozesses“. (2021, Januar). *Diskussionspapier: Modernisierung des Zivilprozesses* (Bericht). Zugriff am 2023-07-03 auf https://www.justiz.bayern.de/media/images/behoerden-und-gerichte/oberlandesgerichte/nuernberg/diskussionspapier_ag_modernisierung.pdf
- Babahmetović, D. & Manresa-Yee, C. (2019, Juni). Towards a Model for Implicit Evaluation in the Wild on a Large Scale. In *Proceedings of the XX International Conference on Human Computer Interaction* (S. 1–4). Donostia Gipuzkoa Spain: ACM. Zugriff auf <https://www.doi.org/10.1145/3335595.3336290>
- Barkhuus, L. & Rode, J. A. (2007, April). From Mice to Men - 24 Years of Evaluation in CHI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. San Jose California USA: ACM. Zugriff auf <https://www.doi.org/10.1145/1240624.2180963>
- Bell, M., Chalmers, M., Barkhuus, L., Hall, M., Sherwood, S., Tennent, P., ... Hampshire, A. (2006, April). Interweaving mobile games with everyday life. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 417–426). Montréal Québec Canada: ACM. Zugriff auf <https://www.doi.org/10.1145/1124772.1124835>

- Benford, S., Greenhalgh, C., Brown, C., Walker, G., Regan, T., Morphett, J., ... Rea, P. (1998, April). Experiments in inhabited TV. In *CHI 98 Conference Summary on Human Factors in Computing Systems* (S. 289–290). Los Angeles California USA: ACM. Zugriff auf <https://www.doi.org/10.1145/286498.286768>
- Benford, S., Seager, W., Flintham, M., Anastasi, R., Rowland, D., Humble, J., ... Sutton, J. (2004). The Error of Our Ways: The Experience of Self-Reported Position in a Location-Based Game. In D. Hutchison et al. (Hrsg.), *UbiComp 2004: Ubiquitous Computing* (Bd. 3205, S. 70–87). Berlin, Heidelberg: Springer Berlin Heidelberg. Zugriff auf https://www.doi.org/10.1007/978-3-540-30119-6_5 (Series Title: Lecture Notes in Computer Science)
- Bergmann, M., Brohmann, B., Hoffmann, E., Loibl, M. C., Rehaag, R., Schramm, E. & Voß, J.-P. (2005). Qualitätskriterien transdisziplinärer Forschung: Ein Leitfaden für die formative Evaluation von Forschungsprojekten. *ISOE-Studentexte* (Nr. 13). Zugriff am 2023-06-20 auf https://www.katalyse.de/wp-content/uploads/2013/08/2005_evalunet_leitfaden.pdf
- Bert, P. (2022a, Oktober). Moderner streiten! Experimente beim digitalen Zivilprozess sind schön und gut – doch es fehlt ein System. *Anwaltsblatt* (10/2022), 529. Zugriff am 2023-04-28 auf <https://anwaltsblatt.anwaltverein.de/files/anwaltsblatt.de/Dokumente/2022/anwbl-2022-529.pdf>
- Bert, P. (2022b, November). Strukturvorgaben im Zivilprozess: Neues aus dem Labor. *Anwaltsblatt. ZPO-Blog*. Zugriff am 2023-08-02 auf <https://anwaltsblatt.anwaltverein.de/de/zpoblog/strukturvorgaben-im-zivilprozess>
- Bert, P. & Windau, B. (2021, Oktober). Reformbedarf im Zivilprozess: Ergebnisse der Umfrage. *Anwaltsblatt. ZPO-Blog*. Zugriff am 2023-04-28 auf <https://anwaltsblatt.anwaltverein.de/de/zpoblog/reformbedarf-zivilprozess-umfrage-ergebnisse>
- Bink, M., Emmert, M., Hellwig, N., Lanzinger, M. & Schönwerth, N. (2022). *Projektbericht zum Praxisseminar 2022 - Implementierung des Basisdokuments*. Zugriff am 2023-06-06 auf <https://github.com/markusbink/basisdokument-implementation/wiki>
- BMI (Hrsg.). (2009). *Arbeitshilfe zur Gesetzesfolgenabschätzung*. Zugriff am 2023-12-19 auf https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/themen/verfassung/arbeitshilfe-gesetzesfolgenabschaetzung.pdf?__blob=publicationFile&v=3
- BMWi (Hrsg.). (2019). *Freiräume für Innovationen – Das Handbuch für Reallabore*. Zugriff am 2023-04-28 auf <https://www.bmwk.de/Redaktion/DE/Publikationen/Digitale-Welt/handbuch-fuer-reallabore.html>
- BMWi (Hrsg.). (2021, September). *Neue Räume, um Innovationen zu erproben*. Zugriff am 2023-05-02 auf <https://www.bmwk.de/Redaktion/DE/Publikationen/Digitale-Welt/konzept-fur-ein-reallabore-gesetz.html>

- Bradley, M. M. & Lang, P. J. (1994, März). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25 (1), 49–59. Zugriff auf [https://www.doi.org/10.1016/0005-7916\(94\)90063-9](https://www.doi.org/10.1016/0005-7916(94)90063-9)
- Brooke, J. (1996). SUS - A quick and dirty usability scale. *Usability evaluation in industry*, 189 (3), 189–194. (Publisher: Taylor & Francis)
- Brown, B., Reeves, S. & Sherwood, S. (2011, Mai). Into the wild: challenges and opportunities for field trial methods. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 1657–1666). Vancouver BC Canada: ACM. Zugriff auf <https://www.doi.org/10.1145/1978942.1979185>
- Butz, A. & Krüger, A. (2017). *Mensch-Maschine-Interaktion* (2., erweiterte auflage Aufl.). Berlin: Walter de Gruyter GmbH.
- Böhm, V., Gebhard, A., Mielke, B. & Wolff, C. (2023). Technikakzeptanz für Legal Tech am Beispiel des Basisdokumentes für den strukturierten Parteivortrag. *Jusletter-IT* (März 2023). Zugriff auf <https://www.doi.org/10.38023/9b534530-73a7-4478-9d74-95d03c4e7f3b>
- Böhret, C. & Konzendorf, G. (2000). *Moderner Staat – Moderne Verwaltung. Leitfaden zur Gesetzesfolgenabschätzung*. Zugriff am 2023-12-19 auf https://www.verwaltung-innovativ.de/SharedDocs/Publikationen/Presse__Archiv/leitfaden_zur_gesetzesfolgenabschaetzung.pdf%3F__blob%3DpublicationFile%26v%3D1
- Chamberlain, A. & Crabtree, A. (Hrsg.). (2020). *Into the Wild: Beyond the Design Research Lab* (Bd. 48). Cham: Springer International Publishing. Zugriff auf <https://www.doi.org/10.1007/978-3-030-18020-1>
- Chesnut, D. & Nichols, K. (2014). *UX for dummies*. Chichester: Wiley.
- Chin, J. P., Diehl, V. A. & Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (S. 213–218). Zugriff am 2023-08-28 auf <https://www.doi.org/10.1145/57167.57203>
- Cockburn, A., Dragicevic, P., Besançon, L. & Gutwin, C. (2020, Juli). Threats of a replication crisis in empirical computer science. *Communications of the ACM*, 63 (8), 70–79. Zugriff auf <https://www.doi.org/10.1145/3360311>
- Crabtree, A. (2004, September). Taking technomethodology seriously: hybrid change in the ethnomethodology–design relationship. *European Journal of Information Systems*, 13 (3), 195–209. Zugriff auf <https://www.doi.org/10.1057/palgrave.ejis.3000500>
- Csikszentmihalyi, M. & Larson, R. (1987). Validity and reliability of the experience-sampling method. *The Journal of nervous and mental disease*, 175 (9), 526–536. (Publisher: LWW)
- Davies, N. (2005). Proof-of-concept demonstrators and other evils of application-led research: A position statement. In *Proceedings of UbiApp Workshop, Munich, Germany*.

- Davis, F. D., Bagozzi, R. P. & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management science*, 35 (8), 982–1003. Zugriff am 2023-09-01 auf <https://www.jstor.org/stable/pdf/2632151.pdf> (Publisher: INFORMS)
- DIN EN ISO 9241-210. (2011). *Ergonomie der Mensch-System-Interaktion – Teil 210: Prozess zur Gestaltung gebrauchstauglicher interaktiver Systeme*.
- Dourish, P., Adler, A., Bellotti, V. & Henderson, A. (1996, März). Your place or mine? Learning from long-term use of Audio-Video communication. *Computer Supported Cooperative Work (CSCW)*, 5 (1), 33–62. Zugriff auf <https://www.doi.org/10.1007/BF00141935>
- Dürr, M., Gröschel, C., Pfeil, U. & Reiterer, H. (2020, April). NurseCare: Design and ‘In-The-Wild’ Evaluation of a Mobile System to Promote the Ergonomic Transfer of Patients. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (S. 1–13). Honolulu HI USA: ACM. Zugriff auf <https://www.doi.org/10.1145/3313831.3376851>
- Echtler, F. & Häußler, M. (2018, April). Open Source, Open Science, and the Replication Crisis in HCI. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (S. 1–8). Montreal QC Canada: ACM. Zugriff auf <https://www.doi.org/10.1145/3170427.3188395>
- Effer-Uhe, D. (2018, Dezember). Strukturierter Parteivortrag im elektronischen Zivilprozess. *Zeitschrift für das gesamte Verfahrensrecht – GVRZ*, 1 (1), 33–46. Zugriff auf <https://www.doi.org/10.9785/gvrz-2018-010106>
- Expertengruppe „Wissenschaft für Nachhaltigkeit“. (2013). *Wissenschaft für Nachhaltigkeit - Herausforderung und Chance für das baden-württembergische Wissenschaftssystem* (Bericht). MWK. Zugriff am 2023-06-20 auf https://mwk.baden-wuerttemberg.de/fileadmin/redaktion/m-mwk/intern/dateien/pdf/Wissenschaft_f%C3%BCr_Nachhaltigkeit/Expertenbericht_RZ_MWK_Broschuere_Nachhaltigkeit_Web.pdf
- Feger, S. S., Dallmeier-Tiessen, S., Woźniak, P. W. & Schmidt, A. (2019, Mai). The Role of HCI in Reproducible Science: Understanding, Supporting and Motivating Core Practices. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (S. 1–6). Glasgow Scotland Uk: ACM. Zugriff auf <https://www.doi.org/10.1145/3290607.3312905>
- Forschungsvertrag. (2022). *Forschungsauftrag „Strukturvorgaben für den Parteivortrag“*. Vertrag zwischen den Justizministerien des Freistaats Bayern und des Landes Niedersachsen und der Universität Regensburg.
- Freisleben, S., Schwarz, S. & Zels, T. (2021). *Projektbericht zum Praxisseminar 2021 (Lehrstuhl Medieninformatik, Universität Regensburg) - Arbeitsgruppe Anwaltschaft*.
- Gaier, R. (2013). Der moderne liberale Zivilprozess. *Neue Juristische Wochenschrift: NJW*, 66 (39), 2871–2876. (Publisher: Verlag CH Beck)
- Galani, A. & Kidd, J. (2019, Februar). Evaluating Digital Cultural Heritage ‘In the Wild’: The Case For Reflexivity. *Journal on Computing and Cultural Heritage*, 12 (1), 1–15. Zugriff auf <https://www.doi.org/10.1145/3287272>

- Gebhard, A. (2022). *Modernisierung des Zivilprozesses - Untersuchung beeinflussender Faktoren bei der Akzeptanz des Basisdokuments* (unveröffentlichte Bachelorthesis). Universität Regensburg, Regensburg.
- Gediga, G., Hamborg, K.-C. & Düntsch, I. (1999, Januar). The IsoMetrics usability inventory: An operationalization of ISO 9241-10 supporting summative and formative evaluation of software systems. *Behaviour & Information Technology*, 18 (3), 151–164. Zugriff auf <https://www.doi.org/10.1080/014492999119057>
- Generosi, A., Ceccacci, S., Faggiano, S., Giraldi, L. & Mengoni, M. (2020). A Toolkit for the Automatic Analysis of Human Behavior in HCI Applications in the Wild. *Advances in Science, Technology and Engineering Systems Journal*, 5 (6), 185–192. Zugriff auf <https://www.doi.org/10.25046/aj050622>
- Geszten, D., Hamornik, B. P. & Hercegi, K. (2019, Oktober). Usability evaluation of a collaborative design software in the wild. In *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)* (S. 101–106). Naples, Italy: IEEE. Zugriff auf <https://www.doi.org/10.1109/CogInfoCom47531.2019.9089963>
- Greger, R. (2019). Der Zivilprozess auf dem Weg in die digitale Sackgasse. *Neue Juristische Wochenschrift (NJW)*, 3429–3432.
- Grunwald, A. (2022). *Technikfolgenabschätzung: Einführung* (3., vollständig aktualisierte und erweiterte Aufl.). Baden-Baden: Nomos.
- Hahn, S., Röhr, I. & Sautmann, M. (2021). *Projektbericht zum Praxisseminar 2021 (Lehrstuhl Medieninformatik, Universität Regensburg) - Arbeitsgruppe Richterschaft*. Zugriff am 2023-06-06 auf <https://github.com/kindOfCurly/PS-Basisdokument/wiki/Projekt-Log>
- Harper, B. D. & Norman, K. L. (1993). Improving user satisfaction: The questionnaire for user interaction satisfaction version 5.5. In *Proceedings of the 1st Annual Mid-Atlantic Human Factors Conference* (Bd. 224, S. 228). Citeseer.
- Harper, R., Lamming, M. & Newman, W. (1992, Dezember). Locating systems at work: implications for the development of active badge applications. *Interacting with Computers*, 4 (3), 343–363. Zugriff auf [https://www.doi.org/10.1016/0953-5438\(92\)90022-8](https://www.doi.org/10.1016/0953-5438(92)90022-8)
- Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology* (Bd. 52, S. 139–183). Elsevier. Zugriff am 2023-03-14 auf [https://www.doi.org/10.1016/S0166-4115\(08\)62386-9](https://www.doi.org/10.1016/S0166-4115(08)62386-9)
- Hassenzahl, M., Burmester, M. & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In G. Szwillus & J. Ziegler (Hrsg.), *Mensch & Computer 2003* (Bd. 57, S. 187–196). Wiesbaden: Vieweg+Teubner Verlag. Zugriff auf https://www.doi.org/10.1007/978-3-322-80058-9_19 (Series Title: Berichte des German Chapter of the ACM)

- Horn, J. (2020, August). Thesenpapier „Modernisierung des Zivilprozesses“ – Auf dem Weg ins 21. Jahrhundert. *Anwaltsblatt. ZPO-Blog*. Zugriff am 2023-04-28 auf <https://anwaltsblatt.anwaltverein.de/de/zpoblog/thesenpapier-modernisierung-des-zivilprozesses-jakob-horn>
- Häußling, R. (2019). *Techniksoziologie: eine Einführung* (2., überarbeitete und aktualisierte Aufl.) (Nr. 4184). Opladen Toronto: Verlag Barbara Budrich.
- IJsselsteijn, W. A., de Kort, Y. & Poels, K. (2013). *The Game Experience Questionnaire*. Technische Universiteit Eindhoven. Zugriff am 2023-09-14 auf <https://research.tue.nl/en/publications/the-game-experience-questionnaire>
- Inan Nur, A., B. Santoso, H. & O. Hadi Putra, P. (2021, Februar). The Method and Metric of User Experience Evaluation: A Systematic Literature Review. In *2021 10th International Conference on Software and Computer Applications* (S. 307–317). Kuala Lumpur Malaysia: ACM. Zugriff auf <https://www.doi.org/10.1145/3457784.3457832>
- Innocenti, P., Hyatt, E. & Harvey, M. (2022, Oktober). Hiker on trail: Information practices in the wild. A mixed methods approach to investigate real-time mobile technology use by long-distance backpackers. *Library & Information Science Research*, 44 (4), 101200. Zugriff auf <https://www.doi.org/10.1016/j.lisr.2022.101200>
- Intille, S. S., Larson, K., Beaudin, J. S., Nawyn, J., Tapia, E. M. & Kaushik, P. (2005, April). A living laboratory for the design and evaluation of ubiquitous computing technologies. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (S. 1941–1944). Portland OR USA: ACM. Zugriff auf <https://www.doi.org/10.1145/1056808.1057062>
- Kilian, M. (2023). Strukturierter Parteivortrag, elektronischer Nachrichtenraum - Was die Anwaltschaft denkt. *Anwaltsblatt* (10/2023), 548–549. Zugriff am 2023-09-26 auf <https://www.anwaltsblatt-datenbank.de/bsab/document/jzs-AnwBl-2023-10-022-548/part/A>
- Kirakowski, J. & Cierlik, B. (1998, Oktober). Measuring the Usability of Web Sites. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 42 (4), 424–428. Zugriff auf <https://www.doi.org/10.1177/154193129804200405>
- Kirakowski, J. & Corbett, M. (1993, September). SUMI: the Software Usability Measurement Inventory. *British Journal of Educational Technology*, 24 (3), 210–212. Zugriff auf <https://www.doi.org/10.1111/j.1467-8535.1993.tb00076.x>
- Kjeldskov, J., Skov, M. B., Als, B. S. & Høegh, R. T. (2004). Is It Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field. In D. Hutchison et al. (Hrsg.), *Mobile Human-Computer Interaction - MobileHCI 2004* (Bd. 3160, S. 61–73). Berlin, Heidelberg: Springer Berlin Heidelberg. Zugriff auf https://www.doi.org/10.1007/978-3-540-28637-0_6

- Koch, K., Tiefenbeck, V., Liu, S., Berger, T., Fleisch, E. & Wortmann, F. (2021, Mai). Taking Mental Health & Well-Being to the Streets: An Exploratory Evaluation of In-Vehicle Interventions in the Wild. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (S. 1–15). Yokohama Japan: ACM. Zugriff auf <https://www.doi.org/10.1145/3411764.3446865>
- Koch, M., Fietkau, J., Draheim, S., Schwarzer, J. & von Luck, K. (2023). Methods and Tools for (Semi-)Automated Evaluation in Long-Term In-the-Wild Deployment Studies. *Mensch und Computer 2023 - Workshopband*. Zugriff auf <https://www.doi.org/10.18420/MUC2023-MCI-WS13-116> (Publisher: GI)
- Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K., Karapanos, E. & Sinnelä, A. (2011, September). UX Curve: A method for evaluating long-term user experience. *Interacting with Computers*, 23 (5), 473–483. Zugriff auf <https://www.doi.org/10.1016/j.intcom.2011.06.005>
- Köbler. (2018). Strukturierter Parteivortrag – das geht doch sowieso nicht. *Deutsche Richterzeitung (DRiZ)*.
- Köbler, R. (2021, Mai). Die kontroverseste Idee: Strukturierter Parteivortrag. Ein Basisdokument als Akte und Tatbestand – der Zivilprozess radikal neu gedacht. *Anwaltsblatt* (05/2021), 283–283. Zugriff am 2023-08-02 auf https://www.anwaltsblatt-datenbank.de/bsab/document/jzs-AnwBl2021050019-000_283/part/A
- Laugwitz, B., Schrepp, M. & Held, T. (2006, Dezember). Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten. In H. Heinecke & H. Paul (Hrsg.), *Mensch und Computer 2006* (S. 125–134). OLDENBOURG WISSENSCHAFTSVERLAG. Zugriff auf <https://www.doi.org/10.1524/9783486841749.125>
- Lavie, T. & Tractinsky, N. (2004, März). Assessing dimensions of perceived visual aesthetics of web sites. *International Journal of Human-Computer Studies*, 60 (3), 269–298. Zugriff auf <https://www.doi.org/10.1016/j.ijhcs.2003.09.002>
- Lazar, J., Feng, J. H. & Hochheiser, H. (2017). *Research methods in human computer interaction* (2. Aufl.). Cambridge, MA: Elsevier.
- Lewis, J. R. (1991, Januar). Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. *ACM SIGCHI Bulletin*, 23 (1), 78–81. Zugriff auf <https://www.doi.org/10.1145/122672.122692>
- Lewis, J. R. (1995, Januar). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7 (1), 57–78. Zugriff auf <https://www.doi.org/10.1080/10447319509526110>
- Liero, K. (2021, März). Digitalisierung des Zivilprozesses: Zuspruch für das Diskussionspapier der Oberlandesgerichte. *Anwaltsblatt*. Zugriff am 2023-04-28 auf <https://anwaltsblatt.anwaltverein.de/de/anwaeltinnen-anwaelte/anwaltspraxis/digitalisierung-des>

-zivilprozesses-zuspruch-fuer-das-diskussionspapier-der-
-oberlandesgerichte

- Lin, H. X., Choong, Y.-Y. & Salvendy, G. (1997, Januar). A proposed index of usability: A method for comparing the relative usability of different software systems. *Behaviour & Information Technology*, 16 (4-5), 267–277. Zugriff auf <https://www.doi.org/10.1080/014492997119833>
- Liu, C., Alavi, H. S., Costanza, E., Zhai, S., Mackay, W. & Moncur, W. (2019, Mai). Rigor, Relevance and Impact: The Tensions and Trade-Offs Between Research in the Lab and in the Wild. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (S. 1–6). Glasgow Scotland Uk: ACM. Zugriff auf <https://www.doi.org/10.1145/3290607.3311744>
- Liu, Y., Dai, W., Feng, C., Wang, W., Yin, G., Zeng, J. & Shan, S. (2022, Oktober). MAFW: A Large-scale, Multi-modal, Compound Affective Database for Dynamic Facial Expression Recognition in the Wild. In *Proceedings of the 30th ACM International Conference on Multimedia* (S. 24–32). Lisboa Portugal: ACM. Zugriff auf <https://www.doi.org/10.1145/3503161.3548190>
- Lund, A. M. (2001). Measuring usability with the USE questionnaire. *Usability interface*, 8 (2), 3–6.
- Maia, C. L. B. & Furtado, E. S. (2016). A Systematic Review About User Experience Evaluation. In A. Marcus (Hrsg.), *Design, User Experience, and Usability: Design Thinking and Methods* (Bd. 9746, S. 445–455). Cham: Springer International Publishing. Zugriff auf https://www.doi.org/10.1007/978-3-319-40409-7_42
- Mamykina, L., Smaldone, A., Bakken, S., Elhadad, N., Mitchell, E., Desai, P., ... Hripcsak, G. (2021, Mai). Scaling Up HCI Research: from Clinical Trials to Deployment in the Wild. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (S. 1–6). Yokohama Japan: ACM. Zugriff auf <https://www.doi.org/10.1145/3411763.3443437>
- Mayring, P. (2010). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (11., aktual. und überarb. Aufl Aufl.). Weinheim Basel: Beltz.
- Mielke, B. & Wolff, C. (2022a). Der strukturierte Parteivortrag im Zivilprozess. *Jusletter-IT* (Mai 2022). Zugriff auf <https://www.doi.org/10.38023/14550731-9195-4c76-a1f7-ba57b5f4914b>
- Mielke, B. & Wolff, C. (2022b). Entwicklung eines digitalen Basisdokuments. *Jusletter-IT* (Mai 2022). Zugriff auf <https://www.doi.org/10.38023/0696d792-d5d8-4367-bc5f-0671da892bf2>
- Namoun, A., Alrehaili, A. & Tufail, A. (2021). A Review of Automated Website Usability Evaluation Tools: Research Issues and Challenges. In M. M. Soares, E. Rosenzweig & A. Marcus (Hrsg.), *Design, User Experience, and Usability: UX Research and Design* (Bd. 12779, S. 292–311). Cham: Springer International Publishing. Zugriff auf https://www.doi.org/10.1007/978-3-030-78221-4_20

- Nielsen, J. (1994, April). *10 Heuristics for User Interface Design*. Zugriff am 2020-02-23 auf <https://www.nngroup.com/articles/ten-usability-heuristics/>
- Nielsen, J. (2001, Januar). *Usability Metrics*. Zugriff am 2023-12-01 auf <https://www.nngroup.com/articles/usability-metrics/>
- Nielsen, J. & Mack, R. L. (Hrsg.). (1994). *Usability inspection methods*. New York: Wiley.
- Norman, D. & Nielsen, J. (o.J.). *The Definition of User Experience (UX)*. Zugriff am 2023-09-14 auf <https://www.nngroup.com/articles/definition-user-experience/>
- Osman Andersen, T., Nunes, F., Wilcox, L., Kaziunas, E., Matthiesen, S. & Magrabi, F. (2021, Mai). Realizing AI in Healthcare: Challenges Appearing in the Wild. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (S. 1–5). Yokohama Japan: ACM. Zugriff auf <https://www.doi.org/10.1145/3411763.3441347>
- Ozok, A. A. & Salvendy, G. (2001, Januar). How consistent is your web design? *Behaviour & Information Technology*, 20 (6), 433–447. Zugriff auf <https://www.doi.org/10.1080/01449290110092260>
- Pettersson, I., Lachner, F., Frison, A.-K., Riener, A. & Butz, A. (2018, April). A Bermuda Triangle?: A Review of Method Application and Triangulation in User Experience Evaluation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (S. 1–16). Montreal QC Canada: ACM. Zugriff auf <https://www.doi.org/10.1145/3173574.3174035>
- Projektgruppe „Reallabore“. (2018, Dezember). *Strategiepapier: Reallabore als Testräume für Innovation und Regulierung* (Bericht). Bundesministerium für Wirtschaft und Energie. Zugriff am 2023-04-28 auf https://www.bmwk.de/Redaktion/DE/Downloads/S-T/strategiepapier-reallabore.pdf?__blob=publicationFile&v=1
- Ranasinghe, C., Pfeiffer, M., Heitmann, S. & Kray, C. (2019, Oktober). Evaluating User Experience under Location Quality Variations: A Framework for in-the-wild Studies. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services* (S. 1–6). Taipei Taiwan: ACM. Zugriff auf <https://www.doi.org/10.1145/3338286.3344392>
- Raven, M. E. & Flanders, A. (1996, Februar). Using contextual inquiry to learn about your audiences. *ACM SIGDOC Asterisk Journal of Computer Documentation*, 20 (1), 1–13. Zugriff auf <https://www.doi.org/10.1145/227614.227615>
- Rogers, Y., Connelly, K., Tedesco, L., Hazlewood, W., Kurtz, A., Hall, R. E., ... Toscos, T. (2007). Why It's Worth the Hassle: The Value of In-Situ Studies When Designing Ubicomp. In D. Hutchison et al. (Hrsg.), *UbiComp 2007: Ubiquitous Computing* (Bd. 4717, S. 336–353). Berlin, Heidelberg: Springer Berlin Heidelberg. Zugriff auf https://www.doi.org/10.1007/978-3-540-74853-3_20

- Rogers, Y. & Marshall, P. (2017). *Research in the Wild*. Cham: Springer International Publishing. Zugriff auf <https://www.doi.org/10.1007/978-3-031-02220-3>
- Root, R. W. (1988). Design of a multi-media vehicle for social browsing. In *Proceedings of the 1988 ACM conference on Computer-supported cooperative work - CSCW '88* (S. 25–38). Portland, Oregon, United States: ACM Press. Zugriff auf <https://www.doi.org/10.1145/62266.62269>
- Rupp, C., Simon, M. & Hocker, F. (2009). Requirements engineering und Management. *HMD Praxis der Wirtschaftsinformatik*, 46 (3), 94–103. (Publisher: Springer)
- Russell, J. A., Weiss, A. & Mendelsohn, G. A. (1989, September). Affect Grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology*, 57 (3), 493–502. Zugriff auf <https://www.doi.org/10.1037/0022-3514.57.3.493>
- Ryan, R. M. (1982, September). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology*, 43 (3), 450–461. Zugriff auf <https://www.doi.org/10.1037/0022-3514.43.3.450>
- Ryan, R. M., Mims, V. & Koestner, R. (1983, Oktober). Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. *Journal of Personality and Social Psychology*, 45 (4), 736–750. Zugriff auf <https://www.doi.org/10.1037/0022-3514.45.4.736>
- Sabanovic, S., Reeder, S. & Kechavarzi, B. (2014, März). Designing Robots in the Wild: In situ Prototype Evaluation for a Break Management Robot. *Journal of Human-Robot Interaction*, 3 (1), 70. Zugriff auf <https://www.doi.org/10.5898/JHRI.3.1.Sabanovic>
- Salovaara, A., Oulasvirta, A. & Jacucci, G. (2017, Mai). Evaluation of Prototypes and the Problem of Possible Futures. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (S. 2064–2077). Denver Colorado USA: ACM. Zugriff auf <https://www.doi.org/10.1145/3025453.3025658>
- Sanchez, C. & Epp, F. A. (2023, März). *Experiential Futures In-the-wild to Inform Policy Design*. arXiv Preprint. Zugriff am 2023-11-20 auf <http://arxiv.org/abs/2303.14174>
- Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. *Appraisal processes in emotion: Theory, methods, research*, 92 (120), 57.
- Schrepp, M., Hinderks, A. & Thomaschewski, J. (2017). Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4 (6), 103. Zugriff am 2023-03-14 auf <https://www.doi.org/10.9781/ijimai.2017.09.001>
- Schwarz, J. R. (1992). *Strukturierter Parteivortrag und elektronische Akte* (Unveröffentlichte Dissertation). Universität Tübingen.

- Schäpke, N., Bergmann, M., Stelzer, F., Lang, D. J. & Guest Editors. (2018, Januar). Labs in the Real World: Advancing Transdisciplinary Research and Sustainability Transformation: Mapping the Field and Emerging Lines of Inquiry. *GAIA - Ecological Perspectives for Science and Society*, 27 (1), 8–11. Zugriff auf <https://www.doi.org/10.14512/gaia.27.S1.4>
- Schäpke, N., Stelzer, F., Bergmann, M., Singer-Brodowski, M., Wanner, M., Caniglia, G. & Lang, D. J. (2017). *Reallabore im Kontext transformativer Forschung: Ansatzpunkte zur Konzeption und Einbettung in den internationalen Forschungsstand* (Bericht Nr. No. 1/2017). Leuphana Universität Lüneburg, Institut für Ethik und Transdisziplinäre Nachhaltigkeitsforschung: IETSR Discussion papers in Transdisciplinary Sustainability Research.
- Shneiderman, B., Plaisant, C., Cohen, M., Jacobs, S. M. & Elmqvist, N. (2017). *Designing the user interface: strategies for effective human-computer interaction* (Sixth Edition Aufl.). Boston: Pearson.
- Sjöman, M. & Hesselgren, M. (2020). Designerly Living Labs: Early-stage exploration of future sustainable concepts. In S. Boess, M. Cheung & R. Cain (Hrsg.), *Proceedings of the Synergy - DRS International Conference 2020*. Zugriff auf <https://www.doi.org/10.21606/drs.2020.307>
- Smith, I., Consolvo, S., Lamarca, A., Hightower, J., Scott, J., Sohn, T., ... Abowd, G. D. (2005). Social Disclosure of Place: From Location Technology to Communication Practices. In D. Hutchison et al. (Hrsg.), *Pervasive Computing* (Bd. 3468, S. 134–151). Berlin, Heidelberg: Springer Berlin Heidelberg. Zugriff auf https://www.doi.org/10.1007/11428572_9
- Tag, B., Goncalves, J., Webber, S., Koval, P. & Kostakos, V. (2022, Januar). A Retrospective and a Look Forward: Lessons Learned From Researching Emotions In-the-Wild. *IEEE Pervasive Computing*, 21 (1), 28–36. Zugriff auf <https://www.doi.org/10.1109/MPRV.2021.3106272>
- Teirlinck, P., Delanghe, H., Padilla, P. & Verbeek, A. (2013, Juni). Closing the policy cycle: Increasing the utilization of evaluation findings in research, technological development and innovation policy design. *Science and Public Policy*, 40 (3), 366–377. Zugriff auf <https://www.doi.org/10.1093/scipol/scs123>
- Thompson, E. R. (2007, März). Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology*, 38 (2), 227–242. Zugriff auf <https://www.doi.org/10.1177/0022022106297301>
- Tolmie, P. & Crabtree, A. (2008, November). Deploying research technology in the home. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work* (S. 639–648). San Diego CA USA: ACM. Zugriff auf <https://www.doi.org/10.1145/1460563.1460662>
- Venkatesh, V. & Bala, H. (2008, Mai). Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences*, 39 (2), 273–315. Zugriff auf <https://www.doi.org/10.1111/j.1540-5915.2008.00192.x>

- Vermeeren, A., Law, E., Roto, V., Obrist, M., Hoonhout, J. & Väänänen-Vainio-Mattila, K. (2010, Oktober). User experience evaluation methods: current state and development needs. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries* (S. 521–530). Reykjavik Iceland: ACM. Zugriff auf <https://www.doi.org/10.1145/1868914.1868973>
- Vorwerk, V. (2017). Strukturiertes Verfahren im Zivilprozess. *Neue Juristische Wochenschrift: NJW*, 70 (32), 2326–2330. (Publisher: Verlag CH Beck)
- Voss, K., Spangenberg, E. & Grohmann, B. (2003, August). Measuring the Hedonic and Utilitarian Dimensions of Consumer Attitude. *Journal of Marketing Research*, 40 (3), 310–320. Zugriff auf <https://www.doi.org/10.1509/jmkr.40.3.310.19238>
- Vries, G. D., Harteveld, M. & Oosterholt, R. (1995, November). Private Camera Conversation: a new method for eliciting user responses? *Behaviour & Information Technology*, 14 (6), 358–360. Zugriff auf <https://www.doi.org/10.1080/01449299508914654>
- Väätäjä, H., Koponen, T. & Roto, V. (2009, Januar). Developing practical tools for user experience evaluation: a case from mobile news journalism. In *Proceedings of the ECCE 2009 - Designing beyond the Product - Understanding Activity and User Experience in Ubiquitous Environments* (S. 8). Zugriff am 2023-09-14 auf https://www.researchgate.net/publication/220956257_Developing_practical_tools_for_user_experience_evaluation_a_case_from_mobile_news_journalism
- Wanner, M., Hilger, A., Westerkowski, J., Rose, M., Stelzer, F. & Schöpke, N. (2018, April). Towards a Cyclical Concept of Real-World Laboratories: A Transdisciplinary Research Practice for Sustainability Transitions. *disP - The Planning Review*, 54 (2), 94–114. Zugriff auf <https://www.doi.org/10.1080/02513625.2018.1487651>
- Watson, D., Clark, L. A. & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, 54 (6), 1063. (Publisher: American Psychological Association)
- Willumeit, H., Gediga, G. & Hamborg, K. C. (1996). IsoMetricsL: Ein Verfahren zur formativen Evaluation von Software nach ISO 9241/10. *Ergonomie und Informatik*, 27, 5–12.
- Witte, J. (2020). Modernisierung des Zivilprozesses. *BRAK Magazin* (5/2020), 4–5.
- Wittig, O. & Schimanek, P. (2020, November). *Reallabore – Überblick über internationale regulatorische Ansätze und ihre Umsetzbarkeit in deutsches Recht* (Gutachten). Ernst & Young Law GmbH für das Bundesministerium für Wirtschaft und Energie. Zugriff am 2023-04-28 auf <https://www.bmwk.de/Redaktion/DE/Publikationen/Studien/reallabore-ueberblick-ueber-internationale-regulatorische-ansaezte.html>

- Youoku, S., Toyoda, Y., Yamamoto, T., Saito, J., Kawamura, R., Mi, X. & Murase, K. (2020, Oktober). *A Multi-term and Multi-task Analyzing Framework for Affective Analysis in-the-wild*. arXiv. Zugriff am 2023-11-20 auf <http://arxiv.org/abs/2009.13885>
- Zwickel, M. (2016). Die Strukturierung von Schriftsätzen. *Monatsschrift für Deutsches Recht*, 70 (16), 988–992. (Publisher: Verlag Dr. Otto Schmidt)
- Zwickel, M. (2023, Februar). Strukturierung des Parteivortrags - Das Spannungsfeld zwischen Parteiherrschaft und Richtermacht. *Anwaltsblatt* (2/2023), 91–93. Zugriff am 2023-04-28 auf <https://anwaltsblatt.anwaltverein.de/files/anwaltsblatt.de/Dokumente/2023/anwbl-2023-091.pdf>

A. Anhang

A.1. Beispiele für standardisierte Fragebögen

Fragebogen	Beschreibung	Quelle
Aesthetics scale	Messung der wahrgenommenen Ästhetik einer Website; fünfstufige Items (Dimensionen: „klassische Ästhetik“ und „expressive Ästhetik“)	Lavie & Tractinsky (2004)
Affect Grid	Messung des Affekts; Ein-Punkt-Skala (Dimensionen: Freude – Unlust, Erregung – Müdigkeit)	Russell et al. (1989)
After Scenario Questionnaire (ASQ)	Messung der Nutzungszufriedenheit bei der Teilnahme an szenariobasierenden Usability-Studien; drei siebenstufige Items (z.B. <i>Overall, I am satisfied with the ease of completing the tasks in this scenario</i>)	Lewis (1991, 1995)
AttrakDiff	Messung wahrgenommener hedonischer und pragmatischer Qualität; 28 Adjektivgegensatzpaare, die auf einer siebenstufigen Skala bewertet werden (z.B. verwirrend – übersichtlich, außergewöhnlich – üblich)	Hassenzahl et al. (2003), https://www.attrakdiff.de
Computer Interface Consistency Questionnaire (ICTQ)	Messung der Auswirkungen der sprachlichen Inkonsistenz des Schnittstellendesigns auf die Nutzendenperformance; 94 neunstufige Items	Ozok & Salvendy (2001)

Computer System Usability Questionnaire (CSUQ) [allgemeinere Version des Post-Study System Usability Questionnaire (PSSUQ)]	Messung wahrgenommener <i>Usability</i> und Zufriedenheit; 19 siebenstufige Items (z.B. <i>I can effectively complete my work using this system, The interface of this system is pleasant</i>)	Lewis (1995)
Geneva Appraisal Questionnaire (GAQ)	Messung der Ergebnisse des Beurteilungsprozesses einer Person, der dem emotionalen Erleben vorausging	Scherer (2001)
Hedonic Utility scale (HED/UT)	Messung hedonischer und utilitaristischer Dimensionen der Verbrauchereinstellung; zehn Adjektivgegensatzpaare (z.B. <i>helpful – unhelpful, dull – exciting</i>)	Voss et al. (2003)
Intrinsic Motivation Inventory (IMI)	Messung der subjektiven Erfahrung der Versuchsteilnehmenden; 45 siebenstufige Items, die je nach Fragestellung unterschiedlich kombiniert werden können (Kategorien: Interesse/Freude, wahrgenommene Kompetenz, Anstrengung, Wert/Nutzen, empfundener Druck, Spannung, wahrgenommene Wahlmöglichkeiten)	Ryan (1982); Ryan et al. (1983)
ISOMetrics	Messung der Nutzungszufriedenheit nach ISO 9241-11; Kurzversion für summative Evaluation mit acht fünfstufigen Items (Kategorien: Eignung für die Aufgabe, Selbstbeschreibungsfähigkeit, Kontrollierbarkeit, Konformität mit den Nutzererwartungen, Fehlertoleranz, Eignung zur Individualisierung, Eignung zum Lernen) und Langversion für formative Evaluation, in der zusätzlich zu jedem Item die Wichtigkeit bewertet wird und Freitext-Feedback möglich ist	Willumeit et al. (1996); Gediga et al. (1999), https://www.isometrics.uni-osnabrueck.de/

NASA Task Load Index (Nasa-TLX)	Messung der wahrgenommenen Arbeitslast; sechs 21-stufige Items (Mentale Anforderung, physische Anforderung, zeitliche Anforderung, Leistung, Anstrengung, Frustration)	Hart & Staveland (1988), https://humansystems.arc.nasa.gov/groups/TLX/
Positive and Negative Affect Schedule (PANAS)	Messung des positiven und negativen Affekts; 20 fünfstufige Items (z.B. <i>Attentive, Irritable, Nervous</i>); inzwischen auch als Kurzversion, übersetzt oder für Kinder verfügbar	Watson et al. (1988); Thompson (2007)
Purdue Usability Testing Questionnaire (PUTQ)	Vergleich der relativen Benutzerfreundlichkeit und Nutzendenperformance verschiedener Softwaresysteme; 100 siebenstufige Items (Kategorien: Kompatibilität, Konsistenz, Flexibilität, Erlernbarkeit, minimale Aktion, minimale Gedächtnisbelastung, Wahrnehmungsbegrenzung, Benutzerführung; z.B. <i>Can users assign command names?, Is the wording consistent with user guidance?</i>)	Lin et al. (1997)
Questionnaire for User Interaction Satisfaction (QUIS)	Messung subjektiver Zufriedenheit der Systemnutzenden; 90 Adjektivgegensatzpaare, die auf einer neunstufigen Skala bewertet werden (bewertete Faktoren in Version 7.0 sind Anzeige, Terminologie und Systemfeedback, Lernen, Systemfunktionen, technische Anleitungen, Online-Tutorials, Multimedia, Telekonferenzen und Softwareinstallation)	Chin et al. (1988); B. D. Harper & Norman (1993)
Self Assessment Manikin (SAM)	Instrument zur Bewertung von Emotionen; grafische Skalen mit drei Gefühlselementen (Kategorien: Vergnügen, Erregung und Dominanz)	Bradley & Lang (1994)

Software Usability Measurement Inventory (SUMI)	Messung der Nutzungsqualität; 50 dreistufige Items (Kategorien: Effizienz, Affekt, Hilfe und Unterstützung, Kontrollierbarkeit und Erlernbarkeit; z.B. <i>This software responds too slowly to inputs, I feel safer if I use only a few familiar functions</i>)	Kirakowski & Corbett (1993), https://sumi.uxp.ie
System Usability Scale (SUS)	Messung der Nutzungsfreundlichkeit; zehn fünfstufige Items (z.B. <i>I found the system unnecessarily complex, I felt very confident using the system</i>)	Brooke (1996)
Technology Acceptance Model (TAM)	Messung der Technologieakzeptanz; neun siebenstufige Items (Kategorien: wahrgenommene Nützlichkeit, wahrgenommene Benutzerfreundlichkeit, Verhaltensabsicht zur Nutzung)	Davis et al. (1989); Venkatesh & Bala (2008)
Usefulness, Satisfaction, and Ease of use Questionnaire (USE)	Messung der Nutzungsfreundlichkeit; 30 siebenstufige Items (Kategorien: Nützlichkeit, Benutzerfreundlichkeit, Leichtigkeit des Lernens, Zufriedenheit; z.B. <i>I would recommend it to a friend, Using it is effortless</i>)	Lund (2001)
User Experience Questionnaire (UEQ)	Messung der Nutzendenerfahrung; 26 Adjektivgegensatzpaare, die auf einer siebenstufigen Skala bewertet werden (Kategorien: Attraktivität, Durchschaubarkeit, Effizienz, Vorhersagbarkeit, Stimulation, Originalität; z.B. <i>creative – dull, usual – leading edge</i>)	Laugwitz et al. (2006); Schrepp et al. (2017)
Website Analysis and Measurement Inventory (WAMMI)	Messung der Attraktivität und Nutzungsfreundlichkeit; 20 fünfstufige Items (Kategorien: Attraktivität, Kontrollierbarkeit, Effizienz, Nützlichkeit und Erlernbarkeit; z.B. <i>This website helps me find what I am looking for, This website is too slow</i>)	Kirakowski & Cierlik (1998), http://www.wammi.com/ , https://wammi.uxp.ie/

Tabelle 1.: Sammlung von Fragebögen, die im UX und *Usability Testing* genutzt werden (alphabetisch sortiert).

A.2. Fragenkatalog für die Evaluation

Im Folgenden ist der **Fragenkatalog** zu finden, der im Rahmen der Evaluation genutzt werden kann bzw. soll. Quellen für diese Fragen sind meine Fokusgruppe zu Projektstand, Projektzielen und Erwartungen an die Evaluation und ein Planungsdokument der Projektgruppe für das Evaluationskonzept. Einzelne Fragen habe ich auf (quantifizierbare) Kriterien heruntergebrochen, für andere scheint das nicht sinnvoll möglich.

Projektziel: Lohnt sich die Einführung des Basisdokuments?

- Personalumschulung notwendig
- kürzere Verfahrensdauer
- kürzere Zeit bis zur mündlichen Verhandlung
- weniger Schriftsatzwechsel
- reduzierte Aufbereitungszeit für Richter:innen vor der Verhandlung
- reduzierter Arbeitsaufwand für Anwält:innen durch Strukturierung
- Einbindung des Basisdokuments in die mündliche Verhandlung
- Einfluss der aktuellen digitalen Ausstattung (Hard- und Software)

Nutzung des Basisdokuments

- Ist das Basisdokument als Werkzeug zur Strukturierung geeignet bzw. wird es so genutzt wie vorgesehen?
 - Nutzung der Bezugnahme
 - Erstellung einer eigenen Relationstabelle nicht nötig
 - reduzierter Umfang des Sachvortrags
 - keine/wenige Redundanzen im Schriftsatz
 - weniger Streitige Punkte
- Wie gut oder schlecht ist der Prototyp nutzbar?⁷⁷
- Eignet sich das Basisdokument für bestimmte Verfahren besser als für andere (z.B. komplex/einfach oder Baukammer/Arzthaftungskammer/...)?
- In welchem Stadium empfiehlt es sich, die Arbeit mit dem Basisdokument zu beginnen (Klageeinreichung/Übergang nach Schriftsatz austausch)?

Umsetzung des Basisdokuments

- Was ist am Basisdokument gut umgesetzt?
 - Schulungen
 - Support

⁷⁷Meiner Einschätzung nach könnte es sinnvoll sein, diese Frage nach der *Usability* des Prototyps in mehrere typische Teilfragen aufzuteilen, z.B. „Was hat für Sie gut funktioniert?“, „Wo sind Sie auf Probleme gestoßen?“ oder „Wie ließ sich der Prototyp in Ihre Arbeitsabläufe integrieren?“

- Onlinehilfe
 - Strukturierungsmöglichkeiten
 - Funktionsumfang
 - Interfacegestaltung
- Was ist am Basisdokument schlecht umgesetzt?
 - Was funktioniert am Prototyp technisch gut?
 - Was funktioniert am Prototyp technisch schlecht?
 - Wie hat sich die Software weiterentwickelt und warum?

Auswirkungen auf das Verfahren

- Wird der Prozess mit Basisdokument für alle Beteiligten übersichtlicher und transparenter?
- Hatte die Arbeit mit dem Basisdokument Auswirkungen auf die Einigungsquote?
- Haben sich durch das optisch dargestellte Gegeneinander im Basisdokument (Relationsaufbau) die Fronten zwischen den Parteien verhärtet?
- Hatte die „gemeinsame“ Arbeit an einem Dokument einen positiven Effekt auf Einigungen untereinander?
- Konnten Rechtsstreitigkeiten frühzeitig in einem Strukturierungstermin geklärt werden? Welchen Einfluss hat das *case management* und die aktive Verfahrensleitung?
- Wie funktioniert die Abschichtung von Parteivortrag?⁷⁸
- Was sind die Auswirkungen der aktiveren Rolle der Richter:innen?
- Erleichtert es das Basisdokument, die Gesamtheit des Vortrags im Blick zu behalten?
- War festzustellen, dass Befangenheitsanträge bei der Arbeit mit dem Basisdokument zunahmen?
- Erleichtert das Basisdokument Sachverständigen die Einarbeitung in den Fall?
- Ändert sich der Gegenstand der mündlichen Verhandlung bei der Arbeit mit dem Basisdokument?
- Ergänzt ein Strukturierungstermin die Arbeit mit dem Basisdokument sinnvoll?
- Eignet sich das Basisdokument als verbindliche Grundlage des Tatbestandes?
- Wäre eine „Verbindlichstellung“ des Sachverhalts nach Bearbeitung des Basisdokuments zu begrüßen?
- Empfiehlt sich für Schriftsätze, die prozessuale Fragen betreffen (z.B. Fristverlängerungen), auch eine Integration in das Basisdokument oder sollten diese weiterhin in herkömmlicher Form erfolgen?

⁷⁸Abschichtung von Parteivortrag bedeutet, dass man den Parteivortrag zeitlich strukturiert, indem man z.B. gestaffelt vorträgt (bzw. vortragen lässt); bspw. tragen die Parteien zuerst zu einem Baumangel A vor, der wird vollumfänglich behandelt, dann zu Mangel B, dann Mangel C, usw.

- Sollten den Parteien innerhalb des Basisdokuments Vorgaben zur Ordnung ihres Vorbringens erteilt werden?
- Wäre eine Präklusion späteren/unstrukturierten Parteivorbringens sinnvoll?

Bedenken bei der Nutzung des Basisdokuments

- Entstanden bei der Arbeit im Basisdokument viele Streitigkeiten über die Struktur an sich?
- Hat sich der/die Beklagte durch Strukturvorgaben seitens der Klägerin/des Klägers erheblich benachteiligt gefühlt?
- Bestehen Schwierigkeiten hinsichtlich der Wahrung richterlicher Neutralität?
- Konnte festgestellt werden, dass Anwält:innen durch die Arbeit im Basisdokument eher geneigt sind, Vorbringen aus prozesstaktischen Gründen zurückzuhalten?
- Wird generell die gestalterische Freiheit der Anwält:innen negativ beeinträchtigt?
- Befürchten Richter:innen Angriffe in einem etwaigen Rechtsmittelverfahren, die mit dem Basisdokument oder Strukturvorgaben zusammenhängen?
- Sind im Zusammenhang mit der Bearbeitung des Basisdokuments Verfahrensfehler zu befürchten?

Arbeitsprozesse der Anwält:innen

- Wie verändert die Arbeit mit dem Basisdokument den anwaltlichen Workflow?
- Wie wirkt sich das Basisdokument auf die Kommunikation zwischen Anwält:innen und Mandant:innen aus?
- Wird die Arbeit mit Textbausteinen beeinflusst?

Arbeitsprozesse der Richter:innen

- Haben Richter:innen durch die Arbeit mit dem Basisdokument vermehrt und punktgenauer richterliche Hinweise erteilt?
- Wurde durch die Arbeit im Basisdokument seltener Vorbringen „übersehen“?
- Bringt das Basisdokument für die Richter:innen Erleichterungen, die über bisher da gewesene Erleichterungen hinaus gehen (z.B. codefy, JuTe)?

A.3. Methodenbeschreibungen

Die Methoden, die in Kapitel 9 diskutiert und in Kapitel 6 noch nicht vorgestellt sind, werden im Folgenden knapp umrissen.

A.3.1. Contextual Inquiry

Forschende führen Vor-Ort-Interviews mit Nutzenden an ihren Arbeitsplätzen durch, während sie arbeiten, um die Struktur der Arbeitspraxis zu beobachten und erfragen. Der:die Forschende agiert als Lehrling, verfolgt, wie Teilnehmende das System nutzen, und fragt, warum. Um die Erfahrung zu untersuchen, achtet man besonders darauf, was positive oder negative Emotionen auslöst, z. B. wie der soziale Kontext die Erfahrung beeinflusst. Drei Aspekte sind zentral für die CI: (a) Die Datenerhebung muss im Kontext der Arbeit Nutzender erfolgen, (b) Forschende und Nutzende bilden eine Partnerschaft, um gemeinsam Fragen zu untersuchen und (c) die Untersuchung basiert auf einem Fokus, d.h. die Untersuchung baut auf einer klar definierten Reihe von Anliegen auf und nicht auf einer Liste spezifischer Fragen.

Für weitere Informationen siehe z.B. Raven & Flanders (1996).

A.3.2. Experience Sampling Method

Methode zur Erfassung von Erfahrungen „in situ“ und unmittelbar, sodass es keine Störungen durch Erinnerungseffekte gibt. Sie kann verwendet werden, um empirische Daten zu erhalten über (a) Häufigkeit und Muster von täglichen Aktivitäten, sozialen Interaktionen und Ortswechseln, (b) Häufigkeit, Intensität und Muster von psychologischen Zuständen (d.h. emotionale, kognitive und konative Dimensionen des Erlebens), (c) Häufigkeit und Muster von Gedanken, einschließlich Qualität und Intensität von Gedankenstörungen. Das Ziel besteht darin, zu ermitteln und zu analysieren, wie die Muster im subjektiven Erleben der Menschen mit den allgemeinen Bedingungen ihres Lebens zusammenhängen. Teilnehmende werden zu unerwarteten, zufälligen Zeiten am Tag dazu aufgefordert, ihre Aktivitäten, Gedanken und inneren Zustand in einem standardisierten Fragebogen festzuhalten.

Für weitere Informationen siehe z.B. Csikszentmihalyi & Larson (1987).

A.3.3. UX Curve

Diese Methode zielt darauf ab, Nutzenden dabei zu helfen, rückblickend zu berichten, wie und warum sich ihre Erfahrungen mit einem Produkt im Laufe der Zeit verändert haben. In der Methode wird eine Vorlage genutzt, mit der die Nutzenden selbst eine Kurve zeichnen können, die beschreibt, wie sich ihre Erfahrungen im Laufe der Zeit entwickelt haben. Die Nutzenden werden dann gebeten, die Gründe an ihren ungefähren Positionen auf der Kurve zu markieren. Die horizontale Achse stellt die zeitliche Dimension vom Beginn der Nutzung bis zum aktuellen Zeitpunkt dar und die vertikale Achse die Intensität der Erfahrungen der Nutzenden. Es gibt vier Kurvenvorlagen, die sich mit UX-spezifischen Gesichtspunkten befassen: der wahrgenommenen Attraktivität, der Benutzerfreundlichkeit, dem Nutzen und dem Grad der Nutzung des Produkts.

Für weitere Informationen siehe Kujala et al. (2011).

A.3.4. Private Camera Conversation

Nach der Nutzung des Systems spricht der:die Teilnehmende mit der Kamera über vorgegebene Themen. So wird Voreingenommenheit des Interviewenden vermieden. Die Videoaufzeichnung kann mehr hedonistische Aspekte zum Vorschein bringen als bei einem Interview, da Teilnehmende dem Interviewenden gegenüber rational handeln wollen. Im Reallabor Strukturvorgaben könnte man die Methode so einsetzen, dass Teilnehmende einen Fragenkatalog bekommen und jeweils nach der Nutzung des Prototyps eine Aufnahme machen.

Für weitere Informationen siehe z.B. Vries et al. (1995).

A.3.5. User Data Logging

In der Anwendung werden Eingaben und Nutzungsverhalten der Nutzenden geloggt und dann ausgewertet. So können Metriken zur Nutzung erhoben werden (Stichwort *Usability Metrics*, siehe z.B. Nielsen (2001)). Im Reallabor Strukturvorgaben müsste dieses Logging innerhalb des Basisdokumentprototyps erfolgen – Nutzendengruppen, die nur mit den Basisdokumentparteien arbeiten, sind in diesen Daten also nicht abgebildet. Bei der Frage, was überhaupt geloggt werden soll, sind Persönlichkeitsrechte⁷⁹ und geltende Datenschutzrichtlinien⁸⁰ zu beachten. So soll das Logging z.B. nicht dazu genutzt werden können, Mitarbeitende zu überwachen. Auch die Schriftsätze unterliegen dem Datenschutz. Bei der Implementierung des Loggings (auch bei der Frage, wie die Daten die Forschenden erreichen und wie sie gespeichert werden dürfen) ist das in jedem Fall genau zu prüfen und transparent zu gestalten.

A.4. Beispiele für Methodensteckbriefe

Diese Methodensteckbriefe sollen als Beispiel zur Veranschaulichung dienen, im digitalen Anhang (Ordner /4_Methodenanalyse) und online⁸¹ sind die Steckbriefe für alle untersuchten Methoden hinterlegt.

A.4.1. Methodensteckbrief: Heuristische Bewertung

Beschreibung: siehe Kapitel 6.8.1.

⁷⁹Persönlichkeitsrecht, Grundgesetz Art. 2 Abs. 1: https://www.gesetze-im-internet.de/gg/art_2.html, abgerufen am 01.12.2023

⁸⁰Datenschutzgrundverordnung (DSGVO): https://www.bmj.de/DE/themen/digitales/DSGVO/DSGVO_node.html, abgerufen am 01.12.2023

⁸¹Link zu den Methodensteckbriefen: https://drive.google.com/file/d/1GIGNktE_OOJV8rGsgea6j-0nrb3POGGi/view?usp=sharing

Rahmenbedingungen Reallabor

- langer Erprobungszeitraum: die heuristische Bewertung wird unabhängig von der Erprobung durchgeführt, kein Einfluss (o)
- lange Verfahren: die heuristische Bewertung wird unabhängig von der Erprobung durchgeführt, kein Einfluss (o)
- unregelmäßige Prototypnutzung: die heuristische Bewertung wird unabhängig von der Erprobung durchgeführt, kein Einfluss (o)
- verschiedene Nutzendengruppen: die heuristische Bewertung wird unabhängig von der Erprobung durchgeführt, kein Einfluss (o)
- Freiwilligkeit: die heuristische Bewertung wird unabhängig von der Erprobung durchgeführt, kein Einfluss (o)
- sensible Daten / Datenschutz: die heuristische Bewertung kann mit Testdaten arbeiten, es ergibt sich keine Problematik (++)
- iterative Entwicklungen am Prototyp: die heuristische Bewertung betrachtet nur eine Version des Prototyps, die Entwicklung wird nicht berücksichtigt (- -)
- im Realbetrieb einsetzbar: die heuristische Bewertung wird unabhängig von der Erprobung durchgeführt und bewertet nicht die Nutzung im Realbetrieb (- -)

Menschliche/organisatorische Aspekte

- geringe Motivation: die Bewertung wird unabhängig von der Erprobung durchgeführt (o)
- schlechte kommunikative Fähigkeiten: die Bewertung wird unabhängig von der Erprobung durchgeführt (o)
- implizites Wissen: die Bewertung wird unabhängig von der Erprobung durchgeführt (o)
- geringes Abstraktionsvermögen: die Bewertung wird unabhängig von der Erprobung durchgeführt (o)
- divergierende Stakeholder-Meinungen: die Bewertung wird unabhängig von der Erprobung durchgeführt (o)
- problematische Gruppendynamik: die Bewertung wird unabhängig von der Erprobung durchgeführt (o)
- hohe Verteilung der Stakeholder: die Bewertung wird unabhängig von der Erprobung durchgeführt (o)
- hohe Komplexität der Systemabläufe: die Bewertung wird unabhängig von der Erprobung durchgeführt (o)
- Meine Erfahrung mit der Ermittlungstechnik: schon heuristische Bewertungen durchgeführt (+)

UX-Metriken

- Aufgabenperformanz: nicht bewertbar, da keine Nutzenden beteiligt sind (- -)
- Zeitperformanz: nicht bewertbar, da keine Nutzenden beteiligt sind (- -)
- Zufriedenheit: nicht bewertbar, da keine Nutzenden beteiligt sind (- -)
- Konsistenz der Anwendung: je nach Heuristik bewertbar (+)
- Ästhetik der Anwendung: je nach Heuristik bewertbar (+)
- Aufgabenlast: nicht bewertbar, da keine Nutzenden beteiligt sind (- -)
- Technologieakzeptanz: nicht bewertbar, da keine Nutzenden beteiligt sind (- -)
- Erwartungen der Teilnehmenden: nicht bewertbar, da keine Nutzenden beteiligt sind (- -)
- Navigation in der Anwendung: je nach Heuristik bewertbar (+)
- Benennungen / Fachbegriffe: nicht bewertbar, da keine Nutzenden beteiligt sind (- -)
- Einfachheit der Nutzung: je nach Heuristik bewertbar (+)
- Sichtbarkeit von Kernfunktionen: je nach Heuristik bewertbar (+)
- Technologieperformanz: je nach Heuristik bewertbar (+)
- Klarheit der Inhalte: je nach Heuristik bewertbar (+)
- Fehleridentifikation: nicht anhand echter Nutzender bewertbar, aber womöglich zu identifizieren (-)
- Fehlerbehandlung: nicht anhand echter Nutzender bewertbar, aber womöglich zu identifizieren (-)
- Zeit zum Lernen der Handhabung: nicht anhand echter Nutzender bewertbar, aber womöglich abschätzbar (-)
- Beibehaltung des Gelernten im Laufe der Zeit: nicht anhand echter Nutzender bewertbar, aber womöglich abschätzbar (-)

Beantwortung der Evaluationsfragen

Die meisten Fragen zur Nutzung des Basisdokument-Prototyps in realen Verfahren können in einer heuristischen Evaluation nicht beantwortet werden. Einzelne Fragen zur Usability können beantwortet werden. (-)

A.4.2. Methodensteckbrief: Interview

Beschreibung: siehe Kapitel 6.6.

Rahmenbedingungen Reallabor

- langer Erprobungszeitraum: Interviews können zu verschiedenen Zeitpunk-

- ten und/oder mehrfach im Erprobungszeitraum genutzt werden, Entwicklungen können so erfasst werden (++)
- lange Verfahren: Interviews können zu verschiedenen Zeitpunkten und/oder mehrfach im Verfahren genutzt werden, Entwicklungen können so erfasst werden (++)
 - unregelmäßige Prototypnutzung: je nach Zeitpunkt des Interviews ist der Prototyp womöglich schon länger nicht mehr genutzt worden (-)
 - verschiedene Nutzendengruppen: verschiedene Interviewleitfäden können für verschiedene Gruppen genutzt werden, um detaillierte Informationen zu sammeln (++)
 - Freiwilligkeit: ein Interview bedeutet Zeitaufwand für Teilnehmende, die Gewinnung könnte sich schwierig gestalten (-)
 - sensible Daten / Datenschutz: im Interview können Teilnehmende selbst beurteilen, welche Daten sie teilen, es ergibt sich also keine Problematik (++)
 - iterative Entwicklungen am Prototyp: ein Interview ist jeweils nur eine Momentaufnahme, unterschiedliche Entwicklungsstadien können durch mehrere Interviews oder gezieltes Fragen evaluiert werden (+)
 - im Realbetrieb einsetzbar: ein Interview kann den realen Kontext einfangen und evaluieren (++)

Menschliche/organisatorische Aspekte

- geringe Motivation: durch die Moderation im Interview kann geringe Motivation der Teilnehmenden halbwegs abgefangen werden, Ergebnisse leiden weniger stark als bei unmoderierten Methoden (+)
- schlechte kommunikative Fähigkeiten: Nachfragen sind während der Sitzung(en) möglich, Unklarheiten können also geklärt werden (++)
- implizites Wissen: implizites Wissen geht verloren, da Teilnehmende keine Abläufe zeigen (- -)
- geringes Abstraktionsvermögen: Nachfragen und Erklärungen sind während der Sitzung(en) möglich, das Problem kann also abgeschwächt werden (+)
- divergierende Stakeholder-Meinungen: im Interview kann jede:r die eigene Meinung darstellen, es ergibt sich keine Problematik (++)
- problematische Gruppendynamik: Teilnehmende werden einzeln befragt, es ergibt sich keine Problematik (++)
- hohe Verteilung der Stakeholder: Interviews können bei zu hohem organisatorischem Aufwand auch remote durchgeführt werden (++)
- hohe Komplexität der Systemabläufe: durch eingehende Befragung können komplexe Abläufe erklärt werden (++)
- Meine Erfahrung mit der Ermittlungstechnik: schon mehrfach Interviews geplant, durchgeführt und ausgewertet (++)

UX-Metriken

- Aufgabenperformanz (- -)
- Zeitperformanz (- -)
- Zufriedenheit: erfragbar (++)
- Konsistenz der Anwendung: erfragbar (++)
- Ästhetik der Anwendung: erfragbar (++)
- Aufgabenlast: nicht messbar, womöglich erfragbar (+)
- Technologieakzeptanz: erfragbar (++)
- Erwartungen der Teilnehmenden: erfragbar (++)
- Navigation in der Anwendung: erfragbar (++)
- Benennungen / Fachbegriffe: erfragbar (++)
- Einfachheit der Nutzung: erfragbar (++)
- Sichtbarkeit von Kernfunktionen: erfragbar (++)
- Technologieperformanz: erfragbar (++)
- Klarheit der Inhalte: erfragbar (++)
- Fehleridentifikation: kann erfragt, aber vermutlich nicht vollständig erfasst werden (+)
- Fehlerbehandlung: kann erfragt, aber vermutlich nicht vollständig erfasst werden (+)
- Zeit zum Lernen der Handhabung: kann erfragt, aber vermutlich nicht vollständig erfasst werden (+)
- Beibehaltung des Gelernten im Laufe der Zeit: kann erfragt, aber vermutlich nicht vollständig erfasst werden (+)

Beantwortung der Evaluationsfragen

Die Fragen können in Interviews mit den verschiedenen Nutzendengruppen beantwortet werden. Einschränkungen bilden die konkreten Erfahrungen der Teilnehmenden: eine kleine Stichprobe hat in den Verfahren womöglich nicht mit allen Aspekten des Basisdokuments gearbeitet bzw. ist damit in Berührung gekommen. (++)

Erklärung zur Urheberschaft

Die vorgelegten Druckexemplare sowie die vorgelegte digitale Version der Arbeit sind identisch.

Ich habe die Arbeit selbständig verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, sowie alle Zitate und Übernahmen von fremden Aussagen kenntlich gemacht.

Ich habe die Arbeit nicht bereits an einer anderen Hochschule zur Erlangung eines akademischen Grades eingereicht.

Von den zu § 27 Abs. 5 der Prüfungsordnung vorgesehenen Rechtsfolgen habe ich Kenntnis.

Regensburg, 31.12.2023

Unterschrift

Erklärung zur Lizenzierung und Publikation dieser Arbeit

Name: Marie Sautmann

Titel der Arbeit: *Evaluation im Reallabor: Analyse und Anwendung von Evaluationsmethoden der Mensch-Maschine-Interaktion in Reallaboren im Rahmen des Forschungsprojekts „Strukturvorgaben für den Parteeivortrag“*

Hiermit gestatte ich die Verwendung der schriftlichen Ausarbeitung zeitlich unbegrenzt und nicht-exklusiv unter folgenden Bedingungen:

- Nur zur Bewertung dieser Arbeit
- Nur innerhalb des Lehrstuhls im Rahmen von Forschung und Lehre
- Unter einer Creative-Commons-Lizenz mit den folgenden Einschränkungen:
 - BY - Namensnennung des Autors
 - NC - Nichtkommerziell
 - SA - Share-Alike, d.h. alle Änderungen müssen unter die gleiche Lizenz gestellt werden.

(An Zitaten und Abbildungen aus fremden Quellen werden keine weiteren Rechte eingeräumt.)

Ich willige ein, dass der Lehrstuhl für Medieninformatik diese Arbeit — falls sie besonders gut ausfällt - auf dem Publikationsserver der Universität Regensburg veröffentlichen lässt.

Ich übertrage deshalb der Universität Regensburg das Recht, die Arbeit elektronisch zu speichern und in Datennetzen öffentlich zugänglich zu machen. Ich übertrage der Universität Regensburg ferner das Recht zur Konvertierung zum Zwecke der Langzeitarchivierung unter Beachtung der Bewahrung des Inhalts (die Originalarchivierung bleibt erhalten).

Ich erkläre außerdem, dass von mir die urheber- und lizenzrechtliche Seite (Copyright) geklärt wurde und Rechte Dritter der Publikation nicht entgegenstehen.

Erklärung zur Lizenzierung und Publikation dieser Arbeit

- Ja, für die komplette Arbeit inklusive Anhang
- Ja, für eine um vertrauliche Informationen gekürzte Variante (auf dem Datenträger beigefügt)
- Nein
- Sperrvermerk bis (Datum):

Regensburg, 31.12.2023

Unterschrift

Inhalt des beigefügten Datenträgers

/1_Ausarbeitung	Die schriftliche Ausarbeitung als PDF und Latex-Projekt
/2_Vorträge	Präsentation des Antrittsvortrags
/3_Fokusgruppe Projektstand	Projektevaluation nach Bergmann et al. (2005): formative Evaluation, Fokusgruppenleitfaden, anonymisiertes Protokoll, Auswertung
/4_Methodenanalyse	Methodensteckbriefe und -matrix
/5_Feldnotizen	Alle Feldnotizen auf die Bezug genommen wird
/6_Bilder	Alle selbst erstellten und aus anderen Quellen übernommenen Bilder