

**Gottfried Wilhelm
Leibniz Universität Hannover
Fakultät für Elektrotechnik und Informatik
Institut für Praktische Informatik
Fachgebiet Software Engineering**

Konzept und Implementierung einer Komponente zur Untersuchung des Erklärungsbedarfs von Software

Bachelorarbeit

im Studiengang Informatik

von

Mathis Fechner

**Prüfer: Prof. Dr. rer. nat. Kurt Schneider
Zweitprüferin: Dr. rer. nat. Jil Ann-Christin
Klünder
Betreuerin: Hannah Luca Deters**

Hannover, 11.05.2023

Erklärung der Selbstständigkeit

Hiermit versichere ich, dass ich die vorliegende Bachelorarbeit selbständig und ohne fremde Hilfe verfasst und keine anderen als die in der Arbeit angegebenen Quellen und Hilfsmittel verwendet habe. Die Arbeit hat in gleicher oder ähnlicher Form noch keinem anderen Prüfungsamt vorgelegen.

Hannover, den 11.05.2023

Mathis Fechner

Zusammenfassung

Software-Systeme sind allgegenwärtig und gewinnen stets an Komplexität. Diese Komplexität kann zu einem Vertrauensverlust des Nutzens in die teils weitreichenden Entscheidungen der Software führen. Darüber hinaus können Bedienungsprobleme oder Ängste in der Bereitstellung von persönlichen Daten auftreten. Ein Weg diese Probleme zu lösen, ist ein Angebot von Erklärungen. Wie viele Erklärungen sinnvoll sind, ist aber von dem Nutzenden abhängig. Die Anzeige von zu vielen Erklärungen kann den Nutzenden durch den kognitiven Aufwand in der Verarbeitung auch einschränken.

Im Rahmen dieser Arbeit wird eine Komponente zur Untersuchung des Erklärungsbedarfs von Software entwickelt. Die Erklärungskomponente stellt auf Nachfrage Erklärungen zu verschiedenen Systemaspekten bereit. Um den Erklärungsbedarf des Nutzens für unterschiedliche Teilszenarien ermitteln zu können, wird die Komponente in einer eigens dafür entwickelten Software Umgebung eingesetzt. In einer Nutzerstudie bearbeiteten 21 Teilnehmende eine Aufgabenstellung in dieser Software Umgebung. Dabei wurde gespeichert, an welchen Stellen zu welchen Systemaspekten Erklärungen angefragt wurden und daraus der Erklärungsbedarf geschlossen. Über das tatsächliche Ziel der Studie wurden die Teilnehmenden erst im Anschluss informiert, um eine unbefangene Softwarenutzung zu sichern.

Die Mehrheit der Teilnehmenden forderte Erklärungen an. Daraus lässt sich ein vermehrt auftretender Erklärungsbedarf schließen. Dieser betraf primär die Bedienung der Software, dessen Umgang mit persönlichen Daten und die inhaltliche Konzeption dessen Features. Eine Komponente, die immer an dem gleichen Ort zu finden ist und auf Nachfrage Erklärungen bereitstellt, wurde von den Teilnehmenden als hilfreich wahrgenommen.

Abstract

Software systems are ubiquitous and are constantly gaining in complexity. This complexity can lead to a loss of trust on the part of the user in the sometimes far-reaching decisions of the software. In addition, operating problems or fears in the provision of personal data can occur. One way to solve these problems is to offer explanations. How many explanations are useful, however, depends on the user. The display of too many explanations can also limit the user by the cognitive effort in processing.

In the context of this thesis a component for the investigation of the explanation need of software is developed. The explanation component provides explanations of various system aspects on demand. In order to be able to determine the user's need for explanations for different sub-scenarios, the component is deployed in a software environment developed specifically for this purpose. In a user study, 21 participants worked on a task in this software environment. In the process, it was recorded at which points explanations were requested for which system aspects, and the need for explanations was concluded from this. The participants were not informed about the actual goal of the study until afterwards in order to ensure unbiased software use.

The majority of participants requested explanations. From this an increased need for explanations can be concluded. This concerned the operation of the software, the handling of personal data and the conception of its features in terms of content. A component that is always in the same place and which provides explanations on request, was perceived as helpful by the participants.

Inhaltsverzeichnis

| | | |
|----------|--|-----------|
| 1 | Einleitung | 1 |
| 1.1 | Problemstellung | 2 |
| 1.2 | Lösungsansatz | 2 |
| 1.3 | Struktur der Arbeit | 3 |
| 2 | Grundlagen | 5 |
| 2.1 | Definition von Erklärbarkeit | 5 |
| 2.1.1 | Algorithmus | 6 |
| 2.1.2 | Privatsphäre | 8 |
| 2.1.3 | Bedienung | 9 |
| 2.2 | Beeinträchtigungen durch Erklärbarkeit | 10 |
| 2.3 | Verwandte Arbeiten | 10 |
| 3 | Forschungsaufbau | 13 |
| 3.1 | Forschungsfragen | 13 |
| 3.2 | Anforderungen an das Forschungsdesign | 14 |
| 3.3 | Forschungsdesign | 16 |
| 3.3.1 | Prototyp | 17 |
| 3.3.2 | Umfrage | 22 |
| 3.4 | Implementierung der Nutzerstudie | 24 |
| 3.4.1 | Verwendete Frameworks | 24 |
| 3.4.2 | Implementierungsdetails | 25 |
| 3.4.3 | Entwicklungsschwierigkeiten | 26 |
| 4 | Auswertung | 27 |
| 5 | Diskussion | 35 |
| 5.1 | Beantwortung von RQ1 | 35 |
| 5.2 | Beantwortung von RQ2 | 36 |
| 5.2.1 | Privatsphäre | 37 |
| 5.2.2 | Algorithmus | 37 |
| 5.3 | Beantwortung von RQ3 | 38 |
| 5.4 | Beantwortung von RQ4 | 39 |
| 5.5 | Limitationen der Studie | 40 |

| | | |
|----------|-------------------------------------|-----------|
| 6 | Zusammenfassung und Ausblick | 41 |
| 6.1 | Zusammenfassung | 41 |
| 6.2 | Ausblick | 42 |
| A | Ein Anhang | 45 |
| A.1 | Umfrage | 45 |

Kapitel 1

Einleitung

Software-Systeme werden zunehmend komplexer und die Aufgabenbereiche, für die softwarebasierte Lösungen gesucht werden, vielseitiger. Während Software zu Beginn der Entwicklung primär eine Unterstützung bei einer Entscheidungsfindung oder redundanter Arbeit darstellen sollte, indem sie klar definierte Berechnungsschritte ausführte, trifft sie die Entscheidungen im Kleinen und Großen zunehmend selbst [15]. So wird beispielsweise entschieden, welche Bewerber es in die nächste Bewerbungsrunde schaffen, welche Route ein Nutzender zum nächsten Termin nimmt und welche Inhalte in dessen News-Feed auf beliebten Social-Media-Plattformen angezeigt werden.

Für genau solche Entscheidungsprobleme gibt es eine stetig wachsende Menge an Lösungen. Viele davon vernachlässigen es dabei aber, Transparenz zu wahren [1]. So wird ein Algorithmus spätestens durch den Einsatz von künstlicher Intelligenz schnell undurchsichtig [11]. Das kann zu diversen Problemen führen. Selbst wenn der Nutzende sich die Funktionsweise einer Software weiterhin erschließen kann, geht schnell das Vertrauen in deren Ergebnisse verloren [6]. Zum Beispiel ist für den Nutzenden nicht ersichtlich, warum ein Navigationssystem eine andere Route empfiehlt, als die dem Nutzenden Bekannte oder auf welcher Grundlage Inhalte auf einer Social-Media-Plattform angezeigt werden. Daraus kann für das erste Beispiel folgen, dass der Nutzende sich mangels Vertrauen für die altbekannte, in dieser Situation aber schlechtere Route entscheidet. Im zweiten Beispiel fühlt der Nutzende sich möglicherweise manipuliert und das Unverständnis über den Algorithmus kann zu einer Unzufriedenheit bis hin zu einem Umstieg auf andere Software führen.

Eine besondere Schwierigkeit ergibt sich dann, wenn zur Ermittlung eines Ergebnisses Daten des Nutzenden erforderlich sind. Datenschutzerklärungen sind meist zu lang und komplex geschrieben, um als Mittel zum Schaffen von Transparenz dienen zu können [13]. Ohne Erklärung fehlt es vielen Nutzenden aber an dem für das Preisgeben von Daten nötigen Vertrauen [5].

Erklärungen können eine Lösung für die genannten Probleme darstellen,

indem sie Nutzenden den Zugang zu den Informationen geben, die sie benötigen, um eine Software soweit zu verstehen und genug Vertrauen aufzubauen, dass sie diese nutzen können [9]. Wie diese Erklärungen aussehen müssen und an welcher Stelle im Entwicklungsprozess von Software Erklärbarkeit als Anforderung berücksichtigt werden muss, ist Teilmenge des Forschungsgebiets der Erklärbarkeit.

1.1 Problemstellung

Wie bereits im vorausgegangenen Kapitel erläutert, kann ohne passende Erklärungen mit wachsender Komplexität von Software-Systemen das Vertrauen der Nutzenden in die Ergebnisse einer Software schwinden oder das Verständnis für die Funktionalität fehlen. In bisherigen Forschungsarbeiten wurde bereits ermittelt, dass Erklärungen gleichzeitig nicht zwangsläufig einen Mehrwert bieten, diese im Gegenteil sogar als Hindernis in der Übersichtlichkeit und Bedienbarkeit einer Software wahrgenommen werden können [3].

Eine wichtige Frage im Forschungsbereich der Erklärbarkeit ist also, in welchen Fällen Erklärungen nötig sind und wie diese aussehen müssen, um eine Hilfe darstellen zu können.

1.2 Lösungsansatz

Wie viele Erklärungen ein Nutzender hilfreich findet und ab wann er sich durch die Menge an Erklärungen beeinträchtigt fühlt, hängt maßgeblich von dem Nutzenden selbst ab. Während ein unerfahrener Nutzender mehr Erklärungen benötigt, stört sich ein Erfahrener an dem höheren Leseaufwand bei der Bedienung. Es ergibt also Sinn, den Nutzenden selbst darüber entscheiden zu lassen, welche Informationen er wann sehen möchte.

In dieser Arbeit wird daher ein Prototyp entwickelt, der eine omnipräsente Erklärungskomponente enthält, über die zu jeder Ansicht Erklärungen abgerufen werden können. Dabei hat der Nutzende die Auswahl zwischen drei Arten von Erklärungen: Erklärungen zum Datenschutz, zum Algorithmus und zur Nutzbarkeit. Hintergründe dazu sind im Grundlagenkapitel 2 zu finden.

Die verfügbaren Erklärungen werden im Rahmen dieser Arbeit statisch hinterlegt und nicht automatisch generiert.

Um ermitteln zu können, wie der Nutzende auf das Erklärungsangebot reagiert, wird in einer Nutzerstudie festgehalten, welche Informationen ein Nutzender wann abrufen. Im Anschluss wird erfragt, wie das Angebot wahrgenommen wurde.

1.3 Struktur der Arbeit

Diese Arbeit ist wie folgt strukturiert. In Kapitel 2 wird die in dieser Arbeit verwendete Definition von Erklärbarkeit vorgestellt. In diesem Schritt werden drei zu erklärende Systemaspekte herausgestellt und erarbeitet, wie diese sich unterscheiden und welche Konsequenz das für die Erklärungen zu den jeweiligen Systemaspekten hat. In Kapitel 3 werden die Forschungsfragen gestellt, ermittelt, welche Anforderungen daraus auf das Forschungsdesign resultieren, das daraus entwickelte Forschungskonzept vorgestellt und auf die Implementierung eingegangen. Im Anschluss werden die gesammelten Forschungsdaten in Kapitel 4 ausgewertet. Die Auswertung wird in Kapitel 5 auf die Forschungsfragen bezogen. Abschließend werden die Ergebnisse dieser Arbeit in Kapitel 6 zusammengefasst und ein Ausblick gegeben.

Kapitel 2

Grundlagen

2.1 Definition von Erklärbarkeit

Forschungsfragen aus dem Bereich der Erklärbarkeit werden schon seit den 1970er Jahren behandelt [7] und vertiefen diverse Aspekte. Es gibt daher unterschiedliche Definitionen davon, was Erklärbarkeit ausmacht [3]. Diese haben zumeist gemeinsam, dass Erklärbarkeit das Bereitstellen von Informationen in Form von Erklärungen erfordert [14].

Lomprozo [12] definiert diese Erklärungen als Prozess und Produkt. Den kognitiven Prozess, in dem die Ursache für etwas ermittelt wird, und die Erklärung selbst als Produkt dieses Prozesses. Miller [14] erweitert diese Definition um einen abschließenden zweiten, den sozialen Prozess. In diesem wird das im ersten Prozess ermittelte Wissen in Form des Produkts, der Erklärung, vermittelt. Die Erklärung wird in dem sozialen Prozess also zu einem Teil einer Konversation. Daraus schließt Miller wiederum, dass Erklärungen kontextabhängig sind. Der Empfangende der Erklärung interessiert sich demnach nur für eine sehr limitierte Menge an Informationen, die in diesem Kontext tatsächlich relevant ist. Welche Informationen das sind, wie das Produkt also aussehen muss und wie diese Informationen in dem kognitiven Prozess ermittelt werden, ist damit abhängig von dem zu erklärenden Element und dem Kontext, in dem ein Nutzender mit diesem konfrontiert wird.

Für Erklärbarkeit im Ganzen wird im Rahmen dieser Arbeit die abstrakte Definition von Chazette et al. [3] verwendet. Chazette et al. betrachten Erklärbarkeit aus der Perspektive des Software-Engineering und definieren sie wie folgt:

Erklärbarkeit 1 *A system S is explainable with respect to an aspect X of S relative to an addressee A in context C if and only if there is an entity E (the explainer) who, by giving a corpus of information I (the explanation of X), enables A to understand X of S in C .*

Die Zuhilfenahme der Formulierung *'in Bezug auf einen Aspekt X von S'* für die Definition von Erklärbarkeit ermöglicht die Verwendung der Definition für unterschiedliche Systemfelder. Dabei wird aber die Frage offen gelassen, in Bezug auf welche Aspekte X eine Software S erklärbar sein muss, um die Software selbst als 'erklärbar' deklarieren zu können. Im Rahmen einer systematischen Literaturrecherche haben Chazette et al. einige zu erklärende Systemaspekte identifiziert und mithilfe von Workshops mit Philosophen und Psychologen verifiziert. Folgende Bereiche waren Ergebnis dieses Prozesses: Das System im Allgemeinen, dessen Argumentationsprozesse, dessen innere Logik, die Interna des Modells und dessen Absichten [3]. Weitere Aspekte, die bei der Literaturrecherche häufig zu finden sind, sind der Umgang des Systems mit persönlichen Daten [2] und die Interaktionsmöglichkeiten des Nutzenden mit dem System [3]. Diese Aspekte ordnen wir in der weiteren Untersuchung in die Bereiche Algorithmus, Privatsphäre und Bedienung ein. In dieser Arbeit wird angenommen, dass diese übergeordneten Systemaspekte grundlegende Aspekte der Erklärbarkeit sind. Deshalb konzentriert sich diese Arbeit auf diese drei Aspekte.

In den folgenden drei Kapiteln werden die drei genannten Aspekte vorgestellt und mit Blick auf den Definitionsansatz von Lomprozo und Miller herausgestellt, welche Informationen der Nutzende im Kontext des entsprechenden Systemaspekts benötigt und wie diese gegebenenfalls ermittelt werden.

2.1.1 Algorithmus

Der Aspekt *Algorithmus* behandelt primär die Argumentationsprozesse, die innere Logik und die Interna des Modells eines Systems. Zusammenfassend also, wie die Ausgaben eines Systems sich aus den Eingaben zusammensetzen. Für den Nutzenden einer Software werden diese Erklärungen in dem Moment interessant, wo ein System eine Ausgabe erzeugt, aber unklar ist, auf welchen Eingaben diese basiert oder zumindest der Zusammenhang zwischen Eingabe und Ausgabe intransparent ist. Ein Beispiel dafür sind automatisch generierte Vorschläge, die in einer Software angezeigt werden [18]. Ein Modell für die Fragen, die sich in diesem Kontext stellen, ist die von Pearl und Mackenzie [16] entwickelte 'Leiter der Kausalität'. In dieser Arbeit wird die von Miller [14] vereinfachte Version verwendet. Diese besteht aus den drei Stufen:

1. Fragen nach dem Subjekt, wie zum Beispiel: Was ist passiert?
2. Modale Fragen, wie zum Beispiel: Wie ist das passiert?
3. Kausale Fragen, wie zum Beispiel: Warum ist das passiert?

1. Fragen nach dem Subjekt

Fragen aus dem ersten genannten Typus muss ein Nutzender sich selbst beantworten, um überhaupt nach Erklärungen fragen zu können. Dafür beobachtet dieser, welche Ereignisse auftreten oder schließt assoziativ aus den beobachteten Ereignissen, welche nicht beobachteten Ereignisse auch aufgetreten sind [14]. Die folgenden Fragetypen bauen auf dem hier gefundenen Ereignis auf.

Für das oben genannte Beispiel – ein Szenario, in dem intransparent generierte Vorschläge in einer Software angezeigt werden – wäre das mit dieser Fragestellung gefundene Ereignis die Einblendung der entsprechenden Vorschläge selbst.

2. Modale Fragen

Die Antworten auf die modalen Fragen geben Erklärungen dafür, wie sich die Ausgabe eines Systems aus der gegebenen Eingabe ergibt. Bei dem oben genannten Beispiel erklären die Antworten auf diesen Fragetypus also, wie und auf welcher Datenbasis die Vorschläge generiert werden. Diese Transparenz ermöglicht es dem Nutzenden, die Software besser zu verstehen, ihr mentales Modell von der Software zu verbessern [14] und den Ergebnissen im Rückschluss mehr zu vertrauen [18]. Damit ist das Erklärungselement, was sich aus den Antworten auf die modalen Fragen ergibt, unter dem Systemaspekt *Algorithmus* für den Nutzenden am wichtigsten.

Sollte es sich bei dem bislang genutzten Beispiel um Vorschläge von Personen auf einer Social-Media-Plattform handeln, könnte die Erklärung hier sein, dass die Person in den eigenen Kontakten ist oder sich die Interessen der vorgeschlagenen Person mit den eigenen decken.

3. Kausale Fragen

Der dritte Fragetypus ist in der Theorie der komplexeste, weil die Beantwortung kontrafaktisches Denken erfordert [14]. Nur durch Simulation alternativer Ursachen kann ermittelt werden, welche Ursachen für das Auftreten des Ereignisses hinreichend verantwortlich sind. Die dadurch gefundene Menge an Kausalketten ist groß und die einzelnen Ketten nehmen darüber hinaus eine immense Länge an. Wie Miller [14] darlegt, denken Nutzende aber kontrastreich. Anstatt zu fragen 'Warum ist das passiert', stellen sie sich daher tatsächlich die Frage, warum eine Option A anstelle einer hypothetischen Option B eingetreten ist. Da Nutzende für den Vergleich zumeist nah aneinander liegende Optionen wählen, unterscheiden sich nur wenige dieser Kausalketten, teilweise jeweils sogar nur in den ersten Elementen. Nutzende sind deshalb gut darin, aus einer großen Menge für den Eintritt eines Ereignisses notwendiger Kausalitäten einige wenige als Erklärung zu akzeptieren.

Handelt es sich bei den Vorschlägen aus dem oben genannten Beispiel um Video Empfehlungen, könnte der Nutzende erwarten, dass wie üblich ein Video eines von ihm favorisierten Inhaltserstellers angezeigt wird. Diese Erwartung entspricht dann einer hypothetischen Option. Hat dieser Inhaltsersteller aber anders als gewöhnlich keine Beiträge produziert, werden andere Videos angezeigt. Die Realität unterscheidet sich dann also, von der hypothetischen Option. Die Differenz liegt in einer Kausalkette, die das Eintreten der realen Option mit einer Änderungen in den verwendeten Daten in Verbindung bringt. Eine Erklärung, die das beschreibt, könnte also sein: 'Da deine favorisierten Inhaltsersteller keine Beiträge bereitgestellt haben, werden Inhalte ähnlicher Produzenten angezeigt.'

Dieser Systemaspekt wirft insbesondere im Kontext von KI die Frage auf, wie die für die Erklärungen relevanten Informationen ermittelt werden können. Da sich diese Arbeit nur mit einer Komponente befasst, die den Erklärungsbedarf des Nutzenden ermitteln soll, ist der Einsatz von statischen, manuell geschriebenen Erklärungen möglich. Die Fragestellung, wie Erklärungen automatisch generiert werden können, wird hier daher nicht behandelt.

2.1.2 Privatsphäre

Während Erklärungen für den *Algorithmus* nach einer Systemausgabe darauf eingehen, wie dieses Ergebnis zustande kam, wird unter diesem Systemaspekt beleuchtet, welche persönlichen Daten durch die Nutzung generiert, gespeichert und preisgegeben werden. Von den Chazette et al. [3] genannten Systemaspekten werden hier primär die innere Logik, die Interna des Modells und dessen Absichten betrachtet. Diese Informationen sollen dabei bereits zugänglich sein, bevor das System persönliche Daten verarbeitet hat. Erklärungen, die diese Informationen bereitstellen, ersetzen nicht die rechtlich bindende Datenschutzerklärung, können aber eine nutzerfreundliche Ergänzung sein und den Nutzenden dazu befähigen, selbst über den Umgang mit den eigenen Daten zu entscheiden [2].

Brunotte et al. [2] beschreiben die Bereitschaft, ein System zu nutzen, als Kompromiss aus den Vorteilen, die es bietet, und der Menge an Daten, die dafür bereitgestellt werden müssen. Beim Schließen dieses Kompromisses muss der Nutzende dem System in dessen Umgang mit den Daten entweder blind vertrauen und seine Privatsphäre damit davon abhängig machen, wie das System arbeitet, oder den *worst case* annehmen und dem System infolge dessen weniger Daten übermitteln. Einem System, das seinen Umgang mit persönlichen Daten erklärt, bringen Nutzende mehr Vertrauen entgegen. [2] Für die Praxis bedeutet das, dass Nutzende einem System eher Daten zur Verfügung, wenn es Informationen darüber anbietet, wie das System die Daten nutzt [5].

Erklärungen im Bereich der *Privatsphäre* sollen den Ergebnissen von Brunotte et al. [2] zufolge drei inhaltliche Bereiche Abdecken:

- **Datennutzung:** Welche Daten werden genutzt? Warum werden sie genutzt? Wie werden sie gesammelt?
- **Datenspeicherung:** Wo und wie lange werden die Daten gespeichert? Wie können die Daten gelöscht werden? Wie werden sie geschützt?
- **Vertraulichkeit:** Wer hat Zugriff auf die Daten? Können die Daten verkauft werden? Werden die Daten mit anderen aggregiert?

In der Form ist zu beachten, dass die Erklärungen gut strukturiert, kurz, präzise und nicht technisch, sondern in einfacher Sprache formuliert sind.

Die größten Ängste im Kontext von Privatsphäre sind bei Nutzenden übermäßige Datenerfassung [8] und daraus folgend Kontrollverlust über die eigenen Daten sowie Datenmissbrauch in Form von invasiver Werbung oder Datendiebstahl [2]. Beispiele für Situationen in denen Erklärungen zu diesem Aspekt hilfreich sind, sind also alle Systeminteraktionen, bei denen Informationen gesammelt werden oder gesammelt werden könnten. Zusammenfassend sind das alle Situationen, in denen ein Nutzender dem System persönliche Daten übermitteln kann oder muss.

2.1.3 Bedienung

Wie Erklärungen zu *Privatsphäre* müssen auch Erklärungen zur *Bedienung* vor dem Eintreten eines Ereignisses bereitgestellt werden. Unter dem Aspekt *Bedienung* wird erklärt, wie ein Nutzender das System verwenden muss, um ein gewünschtes Resultat zu bekommen. Damit kann dieser Systemaspekt als Gegenstück zu dem Aspekt *Algorithmus* gesehen werden. Hier werden also die in 2.1.1 genannten Systemaspekte aus einer anderen Perspektive betrachtet. Anstatt ein gegebenes Ereignis mit Antworten auf modale und kausale Fragen zu erklären, muss hier ein vom Nutzenden potentiell gewünschtes Ereignis und eine Kausalkette, die es eintreten lässt, gefunden und dem Nutzenden bereitgestellt werden. Solche Erklärungen können dabei helfen, Komplexität zu überbrücken [4]. Außerdem können die Erklärungen einem Nutzenden nötige Informationen über die Arbeitsweise eines Systems vermitteln, damit dieser ein mentales Modell über das System aufbauen oder vervollständigen kann, ohne dessen Verhalten für unterschiedliche Eingaben testen zu müssen. Das so aufgebaute Verständnis kann dem Nutzenden dann dabei helfen, bessere Entscheidungen zu treffen [3]. Das ist insbesondere dann sinnvoll, wenn die Software neuartige Bedienmethoden nutzt, für die der Nutzende auch in anderer Software noch kein mentales Modell entwickelt hat.

Für den Systemaspekt Algorithmus wurde als Beispiel genutzt, dass eine Software Vorschläge anzeigt. Der Nutzende könnte einzelne Vorschläge

als unpassend wahrnehmen, diese entsprechend markieren und an deren Stelle einen neuen Vorschlag generieren wollen. Dafür wäre eine Streichgeste möglich, mit der unbeliebte Vorschläge *weggestrichen* werden können. Diese Geste sollte unter dem Systemaspekt *Bedienung* erklärt werden.

2.2 Beeinträchtigungen durch Erklärbarkeit

Erklärbarkeit als Anforderung an ein System bezieht sich nicht auf einzelne Funktionen, sondern auf das System im Ganzen. Damit kann Erklärbarkeit als nicht funktionale Anforderung (NFA) angesehen werden. Wie bei anderen NFAs ist auch die Bewertung von Erklärbarkeit weitgehend subjektiv. Eine Übererfüllung dieser Anforderung oder eine Erfüllung, die Seiteneffekte auf den Rest des Systems ignoriert, kann kontraproduktiv sein.

Erklärbarkeit hat als nicht funktionale Anforderung stets Auswirkungen auf andere nicht funktionale Anforderungen wie Nutzbarkeit [4] und Transparenz [10][3]. Die Unterstützung, die Erklärbarkeit zur Erfüllung dieser Anforderung beitragen kann, wurde im letzten Kapitel schon behandelt. Zu beachten ist aber die kognitive Belastung [14], die zusätzlich bereitgestellte Informationen auf den Nutzenden haben können. Ein Nutzender trifft durch mehr Erklärungen zwar zumeist bessere Entscheidungen, der Zeitaufwand für das Verarbeiten der Informationen kann aber insgesamt die Effizienz einschränken [3]. Zu viele, zu komplexe Informationen können sogar zu einem Gefühl von Kontrollverlust führen [18]. Es ist also unbedingt darauf zu achten, dass Nutzende nur aus ihrer Perspektive notwendige Erklärungen angezeigt bekommen.

Auswirkungen von Erklärbarkeit auf andere Anforderungen zeigen sich schon in der Entwicklung von Software. Dort können Erklärungen hilfreich sein, wenn sie dem Entwickelnden Informationen zur Funktionsweise des Systems vermitteln. Gleichzeitig kann die Implementierung aber auch aufwendig sein und die Instandhaltung des Systems verkomplizieren [3]. In der Konzeption einer Software muss also auch der Entwicklungsaufwand von unterschiedlichen Wegen zur Umsetzung von Erklärbarkeitsanforderungen mit den positiven Einflüssen, die diese Implementierungen auf andere NFAs darstellen, abgewogen werden.

2.3 Verwandte Arbeiten

Miller erforscht in seinem Paper *Explanation in artificial intelligence: Insights from the social sciences* [14], wie eine gute Erklärung auszusehen hat. Dabei stellt er den Nutzenden in den Mittelpunkt und analysiert mithilfe von Forschungsarbeiten aus der Philosophie, Psychologie und den Kognitionswissenschaften, wie Menschen sich gegenseitig Dinge erklären und welche Erklärungen sie am besten verstehen. Dieses Verständnis ist

wichtig für diese Arbeit, damit gegebene Erklärungen gewinnbringend sind und den Nutzenden nicht überlasten. Miller fokussiert sich dabei darauf, wie Erklärungen aussehen und was sie enthalten sollten. Wann Nutzende Erklärungen sehen möchten betrachtet er nur sekundär und nicht im Softwarekontext. Die Frage, wie hoch der Erklärungsbedarf ist, wann er auftritt und wie ein Erklärungsangebot eingebunden werden kann, ohne den Nutzenden zu überfordern, beantwortet er also nicht.

Das Paper *Explaining recommendations in an interactive hybrid social recommender* [18] befasst sich damit, wie von Software generierte Vorschläge besser erklärt werden können. Dabei bezieht es sich ausschließlich auf die modalen Fragen im Systembereich Algorithmus. In dem Paper wurden in einer Nutzerstudie Erklärungen für eine Software getestet, die Wissenschaftlern anhand bestimmter Daten andere Wissenschaftler vorschlägt, denen sie folgen könnten. Neben passiven Erklärungen liegt ein Schwerpunkt dabei auf Anpassungsmöglichkeiten des Algorithmus. Der Nutzende kann die Werte, auf denen die Generierung der Vorschläge basiert, selbst beeinflussen und die Informationsgrundlage dadurch besser verstehen. Diese Methodik ist für das spezifische Einsatzgebiet sehr effektiv, lässt sich aber schon in dem Systembereich Algorithmus schlecht generalisiert einsetzen. In der Nutzerstudie wurden die Regler des interaktive Interfaces viel und die Erklärungen für die dann generierten Vorschläge weit weniger genutzt. Die Vermutung liegt nahe, dass der Erklärungsbedarf durch das interaktive Interface bereits gedeckt ist und die Nachfrage hier nur wenig über das Erklärungsbedürfnis aussagt, was ohne ein komplexes aktives Erklärungselement bestünde. Ein weiteres Ergebnis ist, dass Nutzende die Software dank der Erklärungen zwar als erklärbar einstufen, dafür aber an Kontrollgefühl einbüßen. In den Erklärungen des Papers wird Wert auf Vollständigkeit der Informationen gelegt. Dieses Ziel kann durch die dadurch resultierende Menge an zu verarbeitenden Informationen in Konflikt mit dem Ziel des Nutzenden stehen, die Software bestmöglich zu verstehen. Zu untersuchen ist also, ob auch dann Nachteile für die Erklärungen gesehen werden, wenn nicht das Ziel verfolgt wird, eine Vollständigkeit von Informationen zu garantieren, sondern stattdessen die in Kapitel 2.1.1 erarbeiteten Schwerpunkte berücksichtigt werden. Eine für diese Arbeit wichtige Erkenntnis ist, dass ein Erklärungs-Icon die Bedienbarkeit nicht beeinflusst, dies Format also genutzt werden kann, um ein globales Erklärungsangebot störungsfrei einzubinden.

In der Bachelorarbeit *Developing Personas for Explainability Requirements by Means of a User Study* [17] werden fiktive Identitäten, sogenannte Personas, entwickelt, die Nutzende mit ähnlicher Persönlichkeit und daraus folgendem ähnlichem Erklärungsbedarf gruppieren. Die Personifikationen dieser Gruppen sollen Softwareentwickler dabei unterstützen, Anforderungen an die Erklärbarkeit zu ermitteln. Um die zu gruppierende Personen und ihren Bedarf an Erklärungen ermitteln zu können, wurden eine Umfrage und eine Interviewreihe durchgeführt. Ein wichtiges Ergebnis ist dabei, dass

einige Teilnehmende zwar angeben, Erklärungen sehen zu wollen, sich am Ende aber doch wenig dafür interessieren. Das unterstützt den in dieser Arbeit verfolgten Ansatz, Erklärungen in der praktischen Anwendung auf Nachfrage anzubieten und anhand der Abfragen den Erklärungsbedarf zu ermitteln.

In dem Paper *Privacy explanations - a means to end-user trust* [2] wurde analysiert, wie Nutzende den Erklärungsbedarf von Software unter dem Aspekt Privatsphäre wahrnehmen. Dafür wurden Teilnehmende in einer online-Umfrage mit einer hypothetischen Situation konfrontiert, in der eine Software persönliche Daten abfragt. Dabei zeigte sich, dass viele Teilnehmende zwar Sorge vor übermäßiger Datensammlung und Kontrollverlust über ihre Daten haben, aber trotzdem nur sehr wenige die Datenschutzrichtlinien lesen. Dafür zeigt sich ein Bedarf an einfach zugänglichen Informationen in Form von Erklärungen. Teil der Ergebnisse ist auch, was bei Inhalt und Form der Erklärungen zu beachten ist, wie bereits in Kapitel 2.1.2 aufgegriffen wurde. Wie die im vorherigen Abschnitt vorgestellte Arbeit dargelegt hat, können Nutzende ihren Erklärungsbedarf in der Theorie selbst häufig schlecht einschätzen. Zu untersuchen ist folglich, ob sich dieser Erklärungsbedarf auch in der Praxis zeigt. In dieser Arbeit wird daher getestet, ob Nutzende bei der Abfrage von persönlichen Daten in einer Software also tatsächlich Erklärungen anzeigen lassen.

Das Paper *Explainability as a non-functional requirement: challenges and recommendations* [4] untersucht insbesondere den Zusammenhang zwischen Erklärbarkeit und Transparenz. Dafür wurden Teilnehmende in einer online Umfrage mit der hypothetische Situation konfrontiert, dass ein Navigationssystem eine andere als die gewohnte Route zu einem bestimmten Ziel vorschlägt. In der Umfrage sollten die Teilnehmenden angeben, ob sie sich in einer solchen Situation eine Erklärung für das unerwartete Verhalten wünschen. Die Mehrheit gab dabei an, dass Erklärungen nur auf Anfrage angezeigt werden sollten, was erneut den Ansatz der in dieser Arbeit entwickelten Erklärungskomponente bekräftigt. Eine Erklärungskomponente wurde in dieser Arbeit aber nicht entwickelt und der Erklärungsbedarf auch nicht in der Praxis getestet. Die Teilnehmenden der Studie äußerten sich teilweise positiv über ein Erklärungsangebot, weil es Vertrauen stärkt, die Nutzung vereinfacht und mehr Kontrollgefühl gibt. Andere äußerten wiederum Einschränkungen in der Benutzung durch eine überladene Programmoberfläche, empfanden Erklärungen daher als störend und erwarteten Kontrollverlust durch die Erklärungen. Chazette et al. beschreiben den Effekt von Erklärbarkeit darauf aufbauen als zweischneidiges Schwert, das dem Nutzenden bei richtiger Verwendung dabei hilft, ein System kennenzulernen, effizient damit zu arbeiten und es als transparent wahrzunehmen. Auf der anderen Seite weisen sie darauf hin, dass zu viele, zu kompliziert formulierte Informationen gleichzeitig Unklarheiten auf Seiten des Nutzenden bedeuten und ihn in der Softwarenutzung einschränken können.

Kapitel 3

Forschungsaufbau

3.1 Forschungsfragen

In der Einleitung dieser Arbeit 1 wurde bereits beschrieben, warum Erklärungen in Software wichtig sind. Im vorigen Kapitel 2.1 wurde erarbeitet, dass es verschiedene Systembereiche gibt, für die ein System Erklärungen bereitstellen sollte. Im Anschluss (2.2) wurde aber gezeigt, dass es kontraproduktiv ist, den Nutzenden mit allen verfügbaren Informationen zu den einzelnen Aspekten zu konfrontieren, weil die Erklärungen dann zu einer Überlastung führen. Im Folgenden (2.3) wurde ausgeführt, dass der Erklärungsbedarf des Nutzenden bislang nur weitestgehend theoretisch oder in Form von Nutzerbefragungen zu hypothetischen Situationen erforscht ist. Außerdem wurde festgestellt, dass Nutzende ihren Erklärungsbedarf in für die Nutzerbefragungen verwendeten hypothetischen Situationen selbst unter Umständen falsch einschätzten. Darüberhinaus sind die Forschungsarbeiten bislang auf jeweils einzelne Systemaspekte bezogen. Daraus ergeben sich für diese Arbeit folgende Forschungsfragen:

- RQ1 Wie groß ist der tatsächliche Erklärungsbedarf des Nutzenden in Bezug auf die Systemaspekte Algorithmus, Privatsphäre und Bedienung in der hier verwendeten Beispielsoftware?
- RQ2 In welchen Teilszenarien sind Erklärungen zu welchen Systemaspekten relevant?
- RQ3 Lässt sich der Erklärungsbedarf mithilfe eines omnipräsenten Erklärungsangebots decken, über das auf Nachfrage Erklärungen angezeigt werden?
- RQ4 Werden latente Erklärungen, die erst auf Abfrage bereitgestellt werden, in der Benutzung von Software als hilfreich und vertrauensaufbauend wahrgenommen?

RQ1 fragt rein quantitativ nach dem Erklärungsbedarf von Nutzenden für Software, aufgeschlüsselt auf die drei in Kapitel 2.1 vorgestellten Systemaspekte. RQ2 dient dazu herauszufinden, was dazu führt, dass Nutzende den in RQ1 ermittelten Erklärungsbedarf äußern und entwickelt so einen Zusammenhang zwischen den in Software enthaltenen Teilszenarien und dem in der Benutzung beobachteten Erklärungsbedarf. RQ3 geht nun auf das Format des Erklärungsangebots und die Erklärungen selbst ein und ermittelt, ob das in RQ1 beobachtete Erklärungsbedürfnis, was sich in den in RQ2 aufgeschlüsselten Mustern zeigt, damit decken lässt. In RQ4 wird schließlich betrachtet, ob das reine Angebot von Erklärung in der Benutzung von Software als hilfreich und vertrauensaufbauend wahrgenommen wird, unabhängig davon, ob Nutzende sich die Erklärungen überhaupt anzeigen lassen.

3.2 Anforderungen an das Forschungsdesign

Das Forschungsdesign baut auf verschiedenen Anforderungen auf, die erfüllt sein müssen, um die gestellten Forschungsfragen beantworten zu können. Diese Anforderungen werden im Folgenden mit Bezug auf die entsprechende Forschungsfrage erarbeitet.

Anforderungen durch RQ1

RQ1 erfordert Nutzende als Quelle. Die Beantwortung sollte also auf einer Nutzerstudie aufbauen, in der der Erklärungsbedarf analysiert werden kann. Im Gegensatz zu bestehenden Forschungsarbeiten soll hier nicht die bloße Einschätzung des Nutzenden als Grundlage genutzt, sondern in der Praxis getestet werden, ob der Nutzende einen tatsächlichen Bedarf an Erklärungen zeigt. Dafür soll ein Software-Konzept und ein darauf aufbauender Prototyp entwickelt werden, in dem die Teilnehmenden eine bestimmte Aufgabe erfüllen müssen.

Dieser Prototyp soll in den unterschiedlichen Ansichten über Erklärungen zu den Systemaspekten verfügen, die in der Definition von Erklärbarkeit 2.1 aufgezählten wurden. Um den Nutzenden nicht mit zu vielen Informationen zu überfordern, sollen diese Erklärungen erst auf explizite Aufforderung des Nutzenden hin angezeigt werden. Durch die explizite Aufforderung kann außerdem ermittelt werden, an welcher Stelle tatsächlich eine Erklärung gewünscht war. Um den Erklärungsbedarf auf die verschiedenen Systemaspekte aufschlüsseln zu können, sollte bei Abruf einer Erklärung erfragt werden, für welchen Aspekt gerade eine Erklärung gewünscht ist. Die einzelnen Systemaspekte müssen dem Nutzenden dafür im Vorfeld erklärt werden.

Darüber, dass die Forschungsfragen hinter der Nutzerstudie das Thema Erklärbarkeit behandeln, sollte der Teilnehmende zum Zeitpunkt der Nutzung des Prototyps noch nicht informiert werden. Das sichert die

Unbefangenheit in der Anfrage von Erklärungen und damit valide Forschungsergebnisse.

Die auf Nachfrage gezeigten Erklärungen spielen für die Beantwortung von RQ1 nur eine untergeordnete Rolle. Eine sorgsame Ausarbeitung ist trotzdem gefragt, weil nachlassender Erklärungsbedarf in der Teilnahme an der Studie sonst schlicht auf die Qualität der angezeigten Erklärungen zurückgeführt werden kann.

Anforderungen durch RQ2

Um RQ2 beantworten zu können, müssen bei der Bearbeitung der in dem Prototypen gestellten Aufgabe für jeden in 2.1 genannten Systemaspekt Teilszenarien auftreten, die in den jeweiligen Aspekt einzuordnende Fragen aufwerfen könnten. Solche Teilszenarien können für den Aspekt *Privatsphäre* die Abfrage von persönlichen Daten sein. Für den Aspekt *Algorithmus* sind beispielsweise generierte Vorschläge oder andere auf intransparenten Daten beruhende, komplexe Berechnungen relevant und für den Aspekt *Bedienung* wäre ein solches Teilszenario eine neuartige Bedienungsmethode.

Insbesondere die ersten beiden erwähnten Teilszenarien bedingen einen hohen Grad an Realismus in dem Prototypen, um valide Forschungsergebnisse erzielen zu können. Das unterstützt den Fakt, dass Teilnehmende nicht über das tatsächliche Ziel der Studie informiert werden dürfen. Darüber hinaus erfordert es eine hohe Professionalität in der Implementierung des Prototyps in Bezug auf die Wirkung der Software auf den Teilnehmenden. Nur wenn die Abfragen der Daten echt und die Vorschläge realistisch generiert wirken, lässt sich der sich in solchen Situationen zeigende Erklärungsbedarf untersuchen. Die Teilnahme sollte dafür außerdem in einer natürlichen Umgebung und nicht unter gestellten Laborbedingungen erfolgen.

Um einschätzen zu können, wie Situationen auf einen Teilnehmenden gewirkt haben, ob er beispielsweise von Vorschlägen überrascht war oder die Bedienung als unintuitiv wahrgenommen hat, kann eine Befragung des Nutzenden im Anschluss an die Nutzung des Prototypen hilfreich sein.

Anforderungen durch RQ3

Während für RQ1 primär von Relevanz war, ob überhaupt eine Erklärung angezeigt wurde, spielen bei RQ3 explizit das Format des Angebots und die Erklärung selbst eine Rolle. Um hier valide Ergebnisse zu bekommen, müssen die von dem Teilnehmenden geforderten Erklärungen einen passenden Erklärungsansatz verfolgen. Welche Erklärungsansätze unter welchem Systemaspekt passend sind, wurde bereits in Kapitel 2.1 erarbeitet. Für die Gewinnung der Forschungsdaten muss der Teilnehmende selbst nach seinem Eindruck von der Bedienung des Prototyps gefragt werden. Auch hierfür

sollte nach Abschluss der Aufgabenstellung in dem Prototypen eine Umfrage angeschlossen werden, die Fragen zu der Deckung des Erklärungsbedürfnisses des Teilnehmenden durch die Erklärungen in dem entsprechenden Angebotsformat enthält.

Anforderungen durch RQ4

Bei der vierten Forschungsfrage ist erneut der Eindruck des Teilnehmenden von dem Prototypen bzw. konkret den angezeigten oder zumindest angebotenen Erklärungen nach erfolgreicher Bearbeitung der Aufgabenstellung im Prototypen von Relevanz. Die Umfrage sollte also auch Fragen enthalten, die auf die Beantwortung von RQ4 abzielen. Dafür kann beispielsweise gefragt werden, ob die Erklärungen das Vertrauen in die Software gestärkt haben oder ob Teilnehmende sich ein vergleichbares Angebot in anderer Software wünschen. Im Gegensatz zu RQ3 sind hier auch die Antworten von Teilnehmenden interessant, die keine Erklärungen erfragt haben, weil allein das Angebot schon ein Sicherheitsgefühl auslösen und die Zufriedenheit erhöhen kann.

Weitere Anforderungen

Neben den konkret für eine der Forschungsfragen relevanten Anforderungen ist bei der Umsetzung des Prototypen darauf zu achten, dass die Software übliche usability-Ansätze verfolgt und nicht unnötig umständlich gestaltet ist, um die Forschungsergebnisse nicht mit einem künstlich erzeugten Erklärungsbedarf zu verfälschen.

Die Nutzerstudie sollte neben der Beantwortung von Fragen, die für die einzelnen Forschungsfragen relevant sind, dazu dienen, Grundannahmen zu überprüfen. Es ist beispielsweise möglich, dass Nutzende zwar einen Erklärungsbedarf haben, diesen aber nicht in die drei angebotenen Kategorien einordnen können und infolge dessen trotzdem keine Erklärungen anfordern. Auch eine solche Problemstellung sollte in den Umfrageergebnissen zu erkennen sein. Außerdem sollen hier demographische Daten ermittelt werden, die zur Einordnung der Forschungsdaten relevant sind.

3.3 Forschungsdesign

Der Forschungsaufbau zur Beantwortung der gestellten Forschungsfragen, der sich aus den oben aufgeschlüsselten Anforderungen ergibt, sieht eine Nutzerstudie vor. In dieser Nutzerstudie bearbeiten Teilnehmende zunächst eine Aufgabenstellung in einem eigens entwickelten Prototypen für ein neuartiges Softwarekonzept und reflektieren die Bearbeitung anschließend in einer Umfrage. Im Folgenden werden der Prototyp, die darin umgesetzte Aufgabenstellung und die Umfrage vorgestellt.

3.3.1 Prototyp

In der ersten Stufe der Studie wird dem Nutzenden vermittelt, dass er an einer Studie zum Thema digitale Kommunikation teilnimmt. Dafür wird ein neues Konzept für einen Mikroblogging-Dienst vorgestellt, was der Teilnehmende testen soll. Im Vergleich zu anderen Diensten sollen Beiträge in diesem Konzept um die Möglichkeit erweitert werden, einzelne Textabschnitte mit Referenzen, wie Zeitungsartikeln, zu versehen. Die angelegten Referenzen sollen darüber hinaus mit einem Zitat aus der verlinkten Referenz versehen werden, die einen Zusammenhang zwischen Text und Referenz herstellt. Das kann beispielsweise ein Satz aus einem verlinkten Zeitungsartikel sein, der die eigene Aussage belegt. Die mit Referenz versehenen Textstellen werden in einem Beitrag farblich hinterlegt. Beim darüber Streichen mit der Maus (hovern) wird eine Vorschau der Referenz zusammen mit dem gewählten Zitat angezeigt. Per Klick auf die Referenz kann dann der ganze Kontext des gewählten Zitats gelesen werden. Die genaue Funktionsweise wird weiter unten ausgeführt.

Der für dieses Softwarekonzept entwickelte Prototyp verfügt über ein Erklärungsangebot. Dafür wird unten rechts dauerhaft ein Fragezeichen angezeigt. Beim Streichen über dieses Fragezeichen werden die drei Systemaspekte eingeblendet. Nach der Auswahl wird dann entweder direkt die Erklärung zu dem entsprechenden Systembereich oder eine Übersicht über die unter diesem Aspekt erklärbaren Elemente angezeigt. In dem zweiten Fall wird die Erklärung dann nach der Auswahl des zu erklärenden Elements angezeigt. Welche Erklärungen angezeigt wurden, wird gespeichert. Das ermöglicht das in den für RQ1 gestellten Anforderungen enthaltene Beobachten des Erklärungsbedarfs der Teilnehmenden in der praktischen Anwendung. Andere Daten werden nicht gespeichert. Dem Teilnehmenden wird nicht erklärt, dass dieses Erklärungsangebot Teil der Studie ist. Stattdessen wird darauf hingewiesen, dass der Prototyp über ein Erklärungsangebot verfügt, welches eine Unterstützung in der Arbeit mit der Software bieten soll. Bei der Formulierung der Erklärungen für die einzelnen Ansichten wurde Wert auf zu dem jeweiligen Softwareaspekt passende Erklärungsansätze gelegt. Dafür wurden die in Kapitel 2.1 ausgearbeiteten Anforderungen an Erklärungen herangezogen.

Durch den im Kontext der sozialen Medien neuartigen Ansatz des Softwarekonzepts, Beiträge mit verifizierenden Referenzen zu versehen, und die vollständige Ausrichtung des Prototypen auf dieses Konzept, kann dem Teilnehmenden glaubhaft vermittelt werden, dass er an einer Studie im Bereich digitale Kommunikation teilnimmt. Das stellt die in den Anforderungen von RQ1 enthaltene Unbefangenheit in der Teilnahme sicher. Der Prototyp wurde als Webanwendung, also über einen Browser ausführbares Programm, umgesetzt (siehe 3.4). Das vermittelt, ohne viel Aufwand für den Teilnehmenden, das für die aussagekräftige Beantwortung

von RQ1 und insbesondere RQ2 geforderte Gefühl einer gewöhnlichen Softwarenutzung.

Die Aufgabe des Teilnehmenden ist nun, einen Beitrag für diese Plattform zu schreiben. Diese Aufgabe zeichnet sich in dem für diese Nutzerstudie entwickelten Softwarekonzept durch eine hohe Vielseitigkeit der verschiedenen Szenarien aus. Daher eignet sie sich, um die für RQ2 gestellten Anforderungen zu erfüllen. Im Folgenden wird auf die für das Schreiben des Beitrags nötigen Schritte und deren Relevanz für die tatsächliche Forschungsarbeit eingegangen:

Schritt 0: Einführung in die Studie

Zu Beginn der Studie wird der Teilnehmende zunächst in das Thema eingeführt. Dafür werden in zwei kurzen Sätzen die zu testende Plattform und die Aufgabenstellung erklärt. Ein kurzer Beispielbeitrag visualisiert dem Teilnehmenden das vorgestellte Konzept. Darunter wird das Erklärungsangebot vorgestellt, über das die Software verfügt. Um sicher zu stellen, dass das Erklärungsangebot vollumfänglich wahrgenommen wird, wird der Teilnehmende dazu aufgefordert, alle drei Systemaspekte einmal zu öffnen. Vorher ist das Fortsetzen nicht möglich. Auf dieser einführenden Seite werden in dem Erklärer die einzelnen Softwareaspekte beschrieben, anstatt Erklärungen zur Software selbst anzuzeigen. Diese Beschreibungen lauten wie folgt:

- **Algorithmus:** Hier wird auf kommenden Ansichten erklärt, wie der Algorithmus bei undurchsichtigen Entscheidungen wie zum Beispiel Vorschlägen zu seinen Ergebnissen kommt.
- **Privatsphäre:** Hier wird erklärt, welche Daten auf der aktuellen Ansicht von dir gesammelt werden, und wo, wie lange und zu welchem Zweck sie gespeichert werden. Die Angaben zu Speicherort und Dauer beziehen sich dabei auf die prototypisierte Software.
- **Bedienung:** Hier können Hilfen bei der Bedienung der Software gefunden werden, falls zum Beispiel der nächste Schritt unklar ist oder Fehler auftreten.

Sobald der Teilnehmende sich den Überblick zu allen Systemaspekten anzeigen lassen hat, kann er mit der Bearbeitung der Aufgabenstellung starten.

Schritt 1: Dateneingabe

Der erste Schritt simuliert das Registrieren für die Plattform (Siehe Abb. 3.1). Dafür soll der Teilnehmende Vorname, Nachname, Geburtsdatum und E-Mail-Adresse angeben. Die Eingabe dieser sensiblen persönlicher

Daten ist ein für die Untersuchung des Erklärungsbedarfs des Systemaspekts *Privatsphäre* wichtiges Teilszenario. In der E-Mail-Adresse sind darüber hinaus keine Sonderzeichen erlaubt. Invalide Eingaben führen zu einem Fehler, der über eine rote Umrandung markiert wird. Ein Fehlerfall ist wiederum ein für die Untersuchung des Erklärungsbedarfs des Systemaspekts *Bedienung* interessantes Teilszenario. Bis alle aufgezählten Felder mit validen Daten gefüllt sind, ist der Weiter-Button deaktiviert. Auch das kann zu einem Bedarf an Erklärungen für die *Bedienung* führen.

Registrieren

Bitte trage in dem folgenden Formular die für das Veröffentlichen von Beiträgen notwendigen Daten ein

Vorname

Nachname

Geburtsdatum
04.05.2023

Mailadresse

Weiter

Algorithmus

Privatsphäre

Bedienung

Zurück

Beitrag schreiben

KLIMAKRISE

Die Politik tut zu wenig

Text

0/100-260

Weiter

?

Abbildung 3.1: Ansicht Dateneingabe, ausgeklappter Erklärer

Abbildung 3.2: Beitrag schreiben

Schritt 2: Themenwahl

Nach der Eingabe der persönlichen Daten werden sechs bis sieben mögliche Themenvorschläge angezeigt. Die eine Hälfte der Vorschläge sind dabei zu Themen, die mit Personen aus der entsprechenden Altersgruppe in Verbindung gebracht werden. Die andere Hälfte sind Themen mit Bezug zum aktu-

ellen gesellschaftlichen Diskurs. Mit Blick auf den Systemaspekt *Algorithmus* wird hier erklärt, wie die Vorschläge zustande kommen. Vorschläge sind ein für die Untersuchung des Erklärungsbedarfs dieses Systemaspekts wichtiges Teilszenario. Für den Systemaspekt *Privatsphäre* kann dem Nutzenden wichtig sein, wie mit der Information, welches Thema der Teilnehmende gewählt hat, umgegangen wird.

Schritt 3: Beitrag schreiben

Jetzt wird der Teilnehmende dazu angehalten, einen Beitrag zu dem gerade gewählten Thema zu schreiben (Siehe Abb. 3.2). Der geschriebene Text ist wieder ein persönliches Datum, bei dem der Nutzende unter dem Systemaspekt *Privatsphäre* Fragen zu dem Umgang damit haben könnte. Außerdem gibt es eine Längenvorgabe von 100-260 Zeichen für den zu schreibenden Text. Diese wird unten rechts in dem Textfenster angezeigt. Bis die valide Länge erreicht ist, ist der Weiter-Button deaktiviert. Das kann unter dem Systemaspekt *Bedienung* die Frage aufwerfen, wie man fortsetzen kann.

Schritt 4: Verlinkungen hinzufügen

Nach der Texteingabe soll der Teilnehmende mindestens eine Verlinkung hinzufügen (Siehe Abb. 3.3). Dafür muss er durch Klicken und Streichen mit der Maus, wie beim Markieren von Text, den zu verlinkenden Textbereich auswählen. Für den Systemaspekt *Privatsphäre* ist hier wichtig, wie mit dem ausgewählten Text und den hinzugefügten Referenzen umgegangen wird. Der ausgewählte Text wird im nächsten Schritt an einen Dritten weitergegeben, um Referenzen vorschlagen zu können. Diese Information ist unter dem Aspekt *Privatsphäre* besonders relevant und in der Erklärung enthalten. Das Hinzufügen von Referenzen auf diese Art und Weise ist für die meisten Teilnehmenden vermutlich eine neue Form der Bedienung. Neue Formen der Bedienung wiederum sind ein für die Untersuchung des Erklärungsbedarfs des Systemaspekts *Bedienung* wichtiges Teilszenario. Auch hier ist außerdem der Weiter-Button deaktiviert, bis die erste Referenz hinzugefügt wurde. Erneut kann auch das ein Auslöser für Erklärungsbedarf unter dem Systemaspekt *Bedienung* sein.

Schritt 5: Referenz wählen

Nach der Auswahl des Textes öffnet sich ein Fenster, in dem der ausgewählte Text und ein Textfenster angezeigt werden (Siehe Abb. 3.4). Der Teilnehmende wird nun dazu aufgefordert, eine URL zu einer Referenz zu dem entsprechenden Text einzugeben. Unter dem Systemaspekt *Bedienung* wird an dieser Stelle erklärt, welche Eigenschaften eine URL erfüllen muss, um als valide erkannt zu werden. Solange das Textfeld nicht mit einer



Abbildung 3.3: Verlinkungsansicht mit bereits einer hinzugefügten Verlinkung

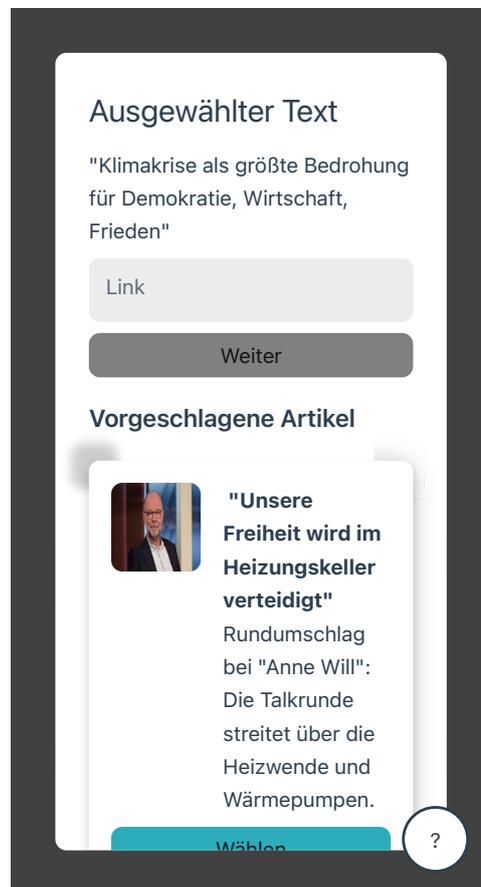


Abbildung 3.4: Referenz finden

validen URL gefüllt wird, ist der Button 'Weiter' deaktiviert. Das kann unter dem Systemaspekt *Bedienung* wieder Erklärungsbedarf verursachen. Unter dem Textfeld werden bereits Vorschläge für Referenzen mit Bild, Titel und einer kurzen Beschreibung angezeigt. Auch diese Vorschläge sind wieder ein für die Untersuchung des Erklärungsbedarfs des Systemaspekts *Algorithmus* wichtiges Teilszenario. Neben den Vorschlägen wird jeweils ein Button 'Wählen' angezeigt, über den das Textfeld für die URL mit der URL der entsprechenden Referenz gefüllt wird. Für den Fall, dass Nutzende sich die Referenz vor dem Auswählen in ganzer Länge ansehen möchten, können sie Bild oder Beschreibung anklicken. Die Referenz wird dann in einem neuen Tab geöffnet. Die Erklärung zum Auswählen und Ansehen der Referenz wird ebenfalls unter dem Systemaspekt *Bedienung* erklärt. Welche Referenz gewählt wird, ist wieder ein persönliches Datum und kann unter dem Systemaspekt *Privatsphäre* eine Erklärung zu dessen Umgang erfordern.

Schritt 6: Zitat hinzufügen

Um die gewählte Referenz mit dem vorher markierten Textabschnitt in Verbindung zu bringen, soll der Teilnehmende nun ein Zitat aus der Referenz in ein Textfeld eingeben, das einen Zusammenhang zu dem gewählten Text herstellt. Unter dem Systemaspekt *Bedienung* wird hier erklärt, dass der Button 'Link hinzufügen' nach Eingabe eines Textes in das Textfeld aktiviert wird. Unter dem Textfeld wird die gewählte Referenz mit Bild, Titel und Beschreibung angezeigt. Wie diese Informationen generiert werden, ist unter dem Systemaspekt *Algorithmus* erklärt. Unter dem Systemaspekt *Privatsphäre* wird erneut beschrieben, wie mit der Information, welches Zitat gewählt wurde, umgegangen wird.

Nach dem Hinzufügen des Zitats wird wieder die aus Schritt Vier bekannte Ansicht angezeigt. Der mit Referenz versehene Textabschnitt ist nun farblich hinterlegt (Siehe 3.3). Neben den bereits in Schritt Vier beschriebenen Erklärungen enthalten die Erklärungen zum Systembereich *Bedienung* noch Informationen darüber, wie bereits erstellte Referenzen gelöscht oder bearbeitet werden können.

Schritt 7: Vorschau

Nach dem Hinzufügen der Referenzen wird eine Vorschau des fertigen Beitrags angezeigt (Siehe 3.5). Der Teilnehmende hat jetzt die Möglichkeit, die angelegten Referenzen durch darüber Streichen mit der Maus auf ihre Korrektheit zu überprüfen (Siehe 3.6). Wurde das noch nicht verstanden, kann es unter dem Systemaspekt *Bedienung* nachgelesen werden. Unter dem Beitrag sind Button zum Veröffentlichen oder Speichern des Beitrags als Entwurf zu finden. Gerade die Möglichkeit des Veröffentlichens ist ein für die Untersuchung des Erklärungsbedarfs des Systemaspekts *Privatsphäre* wichtiges Teilszenario, da es impliziert, dass die zur Verfügung gestellten Informationen für eine größeren Menge an Nutzenden zugänglich gemacht werden und der Teilnehmende somit Kontrolle über die Daten abgeben muss.

Schritt 8: Abschließende Erklärung

Auf der letzten Seite wird der Teilnehmende über das tatsächliche Ziel der Studie aufgeklärt und er bekommt die Möglichkeit, die Teilnahme zu beenden, ohne dass Informationen dazu gespeichert werden. Außerdem wird an dieser Stelle die Umfrage verlinkt, mit der die Teilnahme erfolgreich abgeschlossen werden kann.

3.3.2 Umfrage

Die Umfrage hat primär zum Ziel, die nötige Informationsbasis zur Beantwortung der Forschungsfragen RQ3 und RQ4 aufzubauen. Angaben zu der



Abbildung 3.5: Vorschauansicht Abbildung 3.6: Referenz eingeblendet

Wahrnehmung des Nutzenden helfen aber auch dabei, Ergebnisse für RQ1 und RQ2 zu validieren. Außerdem sollen hier Grundannahmen überprüft und demographische Daten erfasst werden, um die Forschungsergebnisse besser einordnen zu können. Dafür ist die Umfrage in drei Teile aufgeteilt.

Der erste Teil bezieht sich auf den Prototypen und das Erklärungsangebot selbst. Falls ein Teilnehmender keine Erklärungen angefordert hat, wird hier erfragt, ob das tatsächlich darauf zurückzuführen ist, dass kein Bedarf bestand, oder ob das Verständnis für das Angebot oder dessen Kategorien fehlte. So können Ergebnisse aus dem Prototypen für RQ1 validiert werden. Außerdem kann der Teilnehmende hier angeben, ob das Erklärungsbedürfnis gedeckt werden konnte, und auf offene Fragestellungen hinweisen, die der Teilnehmende nicht in die gegebenen Kategorien einordnen konnte. Darüber hinaus wird erfragt, ob das Angebot als hilfreich oder sogar störend wahrgenommen wurde und ggf. darum gebeten, die Ursache dafür zu nennen. Diese Informationen sind zur Beantwortung der Forschungsfragen RQ3 und RQ4 wichtig.

Der zweite Teil behandelt die drei Systemaspekte. Für jeden dieser Aspekte gibt es eine Unterkategorie, in der Teilnehmende auf einer fünfstufigen Likert-Skala ihre Zustimmung zu bis zu sechs Aussagen angeben sollen. Anhand dieser Aussagen kann ermittelt werden, ob die im Prototypen aufgebauten Teilszenarien so auf den Nutzenden gewirkt haben, wie erwartet. Dadurch braucht bei der Auswertung der Studie nicht auf Basis der Grundannahmen gearbeitet zu werden, die zur Entwicklung der Teilszenarien herangezogen wurden. Beispielsweise ist es möglich, dass die persönlichen Daten im Rahmen der Studie nicht als schützenswert angesehen wurden. Unter dem Aspekt *Algorithmus* könnten Vorschläge als so gewöhnliche Softwareeigenschaft bewertet werden, dass sie keines weiteren Hinterfragens bedürfen, oder ein Teilnehmender hat bereits selbst eine vermeintliche Ursache ausgemacht. Unter dem Aspekt *Bedienung* könnte die Bedienung von als neuartig konzipierten Elementen als intuitiv wahrgenommen werden, in dem aus anderen mentalen Modellen Wissen transferiert wird. Diese Metadaten zu den über den Prototypen ermittelten Daten unterstützen die Beantwortung von RQ2. Für die Beantwortung von RQ3 und RQ4 enthalten die Kategorien darüber hinaus Aussagen zur Zufriedenheit mit den Erklärungen und, unabhängig von der tatsächlichen Anforderung, Aussagen dazu, ob die Erklärungen das Vertrauen in die Software gestärkt haben.

Im letzten Teil werden demographische Daten der Studienteilnehmenden ermittelt. Diese enthalten das Alter, das Geschlecht und den beruflichen Hintergrund. Außerdem werden in diesem Teil Fragen gestellt, die auf die technischen Fähigkeiten der Teilnehmenden schließen lassen.

3.4 Implementierung der Nutzerstudie

3.4.1 Verwendete Frameworks

Eine Anforderung an die Implementierung des Prototypen ist, dass er von Teilnehmenden ohne großen Aufwand in einer gewohnten Umgebung ausgeführt werden kann. Wie bereits erwähnt, wird deshalb eine sogenannte Webanwendung entwickelt. Eine Webanwendung unterscheidet sich von einer gewöhnlichen Desktop-Anwendung darin, dass sie nicht installiert zu werden braucht, sondern in einem Browser ausgeführt werden kann. Genauer handelt es sich in dieser Arbeit um eine Einseiten-Webanwendung. Im Gegensatz zu gewöhnlichen Webseiten und anderen Webanwendungen werden diese nur einmalig geladen. Anstatt bei Interaktion des Nutzenden auf eine neue Seite zu verlinken, laden diese weitere Inhalte in der Nutzung dynamisch nach oder generieren die Inhalte selbst. Diese Entwicklungsform ist inzwischen weit verbreitet und viele Arbeitsschritte werden von Javascript Frameworks übernommen oder zumindest vereinfacht. Das hier genutzte und auch darüber hinaus sehr beliebte Framework ist 'Vue.js'.

Vue übernimmt das Zusammensetzen von einzelnen Komponenten zu

einer Ansicht. Komponenten können dabei Button oder auch größere Kompositionen wie ein Beitrag mit Text und Verlinkungen sein. Mit dem 'Vue Router' können mehrere Einzelansichten strukturiert werden. Dafür wird in der Hauptansicht eine Komponente eingebaut, die der Router mit der aktuell anzuzeigenden Ansicht ersetzt. Darüber hinaus bietet vue einfache Javascript Schnittstellen, um HTML und CSS, also die Struktur und das Aussehen der Webseite, zu manipulieren und auf Interaktionen des Nutzenden zu reagieren.

Im Backend brauchen nur noch wenige Schritte, wie das Ermitteln der Vorschau-Daten für die Referenzen, die Referenzen selbst und das Speichern der forschungsrelevanten Daten umgesetzt zu werden. Dafür wurde das Python basierte Framework 'Flask' verwendet. Die anschließende Studie wurde mit 'LimeSurvey' umgesetzt.

3.4.2 Implementierungsdetails

Die Anzeige für die Erklärungen ist in der Hauptansicht untergebracht und wird so über allen vom Router eingesetzten Teilansichten angezeigt. Das bedeutet, dass der Erklärer in den einzelnen Unteransichten nicht neu eingebunden zu werden braucht. Die verfügbaren Erklärungsobjekte werden in einem Javascript-Objekt gesammelt und über einen Routen-Namen gefunden. Der Erklärer ermittelt also über den Vue-Router, welchen Namen die aktuelle Ansicht hat, und zeigt dann die im Javascript-Objekt zu diesem Namen gespeicherten Erklärungen an.

Die Erklärungsobjekte enthalten drei Eigenschaften für Erklärungen zu den drei Systemaspekten. Die Eigenschaften können dabei leer sein, falls es auf der aktuellen Ansicht keine Erklärung zu dem entsprechenden Systembereich gibt, einen String für die Erklärung selbst oder ein weiteres Objekt enthalten. Über das weitere Objekt ist die Strukturierung von Erklärungen in einzelne Bereiche, wie einem bestimmten zu erklärenden Button und einem Texteingabefeld, möglich. So können für die verschiedenen Anforderungen der unterschiedlichen Ansichten mit geringem Aufwand Erklärungen angelegt werden.

Dank dieser Implementierung brauchen neue Ansichten nur in dem Router angelegt, mit einem Namen versehen und in dem Javascript-Objekt die entsprechenden Erklärungen ergänzt zu werden. Das bedeutet für die Entwickler einen sehr geringen Mehraufwand durch die Erklärungen und reduziert so die in Kapitel 2.2 vorgestellten Entwicklungseinschränkungen.

Um zu speichern, welche Erklärungen angezeigt werden, wird in dem Prototypen aus dem Javascript-Objekt mit den Erklärungen ein zweites Objekt mit gleicher Struktur generiert. In diesem 'Rückmeldungs'-Objekt wird statt der tatsächlichen Erklärungen ein Boolean gespeichert, der beim Ansehen einer Erklärung auf wahr gesetzt wird.

Auf der letzten Ansicht wird das 'Rückmeldungs'-Objekt an den Server gesendet. Auf dem Server wird das Objekt dann für die weitere Verarbeitung

gespeichert und ein noch unbenutzter Schlüssel für die LimeSurvey Umfrage herausgesucht. Dieser Schlüssel wird ebenfalls gespeichert und als Antwort auf die Anfrage an den Server übermittelt. Der Prototyp bettet den Schlüssel dann in einen Link ein, über den der Teilnehmende an die Umfrage weitervermittelt wird. So können die Ergebnisse aus dem Prototypen mit denen aus der Umfrage verknüpft werden.

3.4.3 Entwicklungsschwierigkeiten

Die größte Schwierigkeit in der Entwicklung lag in der Generierung von Vorschlägen. Gibt der Teilnehmende selbst eine URL für einen Artikel ein, können Open-Graph-Informationen genutzt werden, um die Vorschau zu erhalten. Open-Graph-Informationen sind Tags, die im HTML einer Webseite gesetzt werden und einen Titel, einen Beschreibungstext und einen Link zu einem Bild enthalten können. Um URLs für ausgewählte Textstellen überhaupt erst mal zu finden, sollte eine News API genutzt werden. Wenn beim Suchen nach dem Text keine Ergebnisse gefunden wurden, wurden alle kleingeschriebenen Worte aus dem Text entfernt, um dadurch Schlagworte für die Suche zu generieren. Brachte auch das keine Ergebnisse, wurde die Liste an Schlagworten sukzessive verkürzt. Dadurch entstehen bei einer Suche unter Umständen mehrere Anfragen an den Server. Bei einer direkten Anfrage an die Google News Server wurde die eigene IP Adresse bereits nach wenigen Versuchen blockiert. Auch eine anfragenärmere Implementierung hatte eine schnelle Blockierung durch Google zur Folge. Als Lösung wurde schließlich die API 'NewsApi' verwendet, die zwar ein tägliches Limit von 100 Anfragen vorsieht, bis zu dessen Überschreitung aber ausreichend schnelle Antworten liefert.

Kapitel 4

Auswertung

Die bei der Durchführung der im letzten Kapitel vorgestellten Nutzerstudie gesammelten Daten werden im Folgenden ausgewertet.

Wie bereits in 3.3.2 beschrieben, wurden viele der genutzten Daten über Likert Skalen ermittelt. Vorweg sei angemerkt, dass es sich bei Daten aus einer Likert Skala zumeist um ordinal verteilte Punkte handelt. Das bedeutet, dass die für den Teilnehmenden zur Auswahl stehenden Werte diskret und ihre Abstände nicht quantifizierbar sind. Bei der Auswertung dieser Umfrage nehmen wir an, dass die Likert Skala symmetrisch formuliert wurde, gehen also davon aus, dass die Skalenpunkte von allen Befragten als äquidistant wahrgenommen wurden. So können die Daten als Datenpunkte einer metrischen Skala interpretiert werden, was Auswertungswerkzeuge wie den Mittelwert und die Standardabweichung zulässt.

Die Likert Skalen hatten die fünf Stufen 'Stimme nicht zu', 'Stimme eher nicht zu', 'neutral', 'Stimme eher zu' und 'Stimme zu'. Für die Berechnungen wurden diese Stufen auf die Skala -2 bis +2 abgebildet. Ein Wert von -2 entspricht einer entschiedenen Ablehnung einer Aussage und ein Wert von 2 einer ebenso entschiedenen Zustimmung. Zudem gab es die Möglichkeit, keine Aussage zu tätigen.

Die Auswertung ist folgendermaßen strukturiert: Zu Beginn werden die demographischen Daten betrachtet, um im Folgenden Bezug darauf nehmen zu können. Im Anschluss wird ausgewertet, in welchen Teilszenarien im Prototypen wie viele Erklärungen zu welchen Systemaspekten aufgerufen wurden. Diese Ergebnisse werden dann mithilfe der Umfrageergebnisse zu dem Prototypen im Allgemeinen in einen Kontext gesetzt. Zuletzt werden die Umfrageergebnisse zu den einzelnen Systemaspekten analysiert.

Demographische Daten (Teil drei der Umfrage)

Die Studie hat insgesamt 21 Teilnehmende. 67% dieser Teilnehmenden waren männlich. Die älteste teilnehmende Person war 62 Jahre, die jüngste 20 Jahre alt. Das Durchschnittsalter lag bei 27 Jahren. Wie am Durchschnittsalter

schon zu erkennen ist, sind die Altersgruppen unterschiedlich stark vertreten. So waren 86% der Teilnehmenden unter 30, 67% sogar unter 25 Jahren alt. Die Hälfte der Teilnehmenden arbeitet oder studiert im technischen Bereich. Bei der anderen Hälfte ist der berufliche Hintergrund über den handwerklichen, medizinischen, sozialen und wirtschaftlichen Bereich verstreut. Zwei Teilnehmende machten an dieser Stelle keine Aussage.

Um nicht allein aus Alter und beruflichem Hintergrund auf die technischen Fähigkeiten schließen zu müssen, wurde in zwei weiteren Fragen ermittelt, ob Teilnehmende viel mit Softwaresystemen arbeiten und wie sicher sie sich in der Bedienung von Software fühlen (Siehe Anhang A.1: Umfrage). Aus den zwei Angaben wurde der Mittelwert gebildet und aus diesem Mittelwert geschlossen, ob eine Person als technisch eher versiert oder eher unversiert anzusehen ist. Demzufolge sind 71% der Teilnehmenden als eher technisch versiert einzustufen. Entgegen der Erwartung haben sich 75% der über 25 jährigen und nur 64% der unter 25 jährigen Teilnehmenden als technisch versiert eingestuft.

Anzeige von Erklärungen (Prototyp)

Insgesamt 76% der Teilnehmenden haben sich mindestens eine Erklärung anzeigen lassen. Die übrigen 24% der Teilnehmenden gaben an, keine Erklärungen angefragt zu haben, weil sie keine Fragen hatten, für die sie gerne eine Erklärung gehabt hätten. Die Teilnehmenden, die Erklärungen anfragten, forderten durchschnittlich 6,1 Erklärungen an. Eine genauere Aufschlüsselung auf die Teilnehmenden zeigt folgende Tabelle:

| | min. 25 Jahre | jünger als 25 | technisch eher versiert | technisch eher unversiert |
|--|---------------|---------------|-------------------------|---------------------------|
| Erklärungen angezeigt in % | 100% | 64% | 73 % | 83% |
| Durchschnittliche Anzahl angezeigter Erklärungen | 7,3 | 5,1 | 6,7 | 5,4 |

Alle Personen, die sich Erklärungen haben anzeigen lassen, liessen sich mindestens eine und im Durchschnitt 3,1 Erklärungen aus dem Systembereich Bedienung anzeigen. Im Bereich Algorithmus und Privatsphäre öffneten 57% der Teilnehmenden mindestens eine und durchschnittlich 1,8 Erklärungen.

Die meisten Erklärungen wurden in den folgenden Ansichten angezeigt:

| Ansicht 3.3.1 | Systemaspekt | Anzahl aufgerufener Erklärungen |
|-------------------------------------|--------------|---------------------------------|
| Verlinkungen hinzufügen (Schritt 4) | Bedienung | 10 |
| Dateneingabe (Schritt 1) | Privatsphäre | 8 |
| Beitrag schreiben (Schritt 3) | Bedienung | 8 |
| Beitrag schreiben (Schritt 3) | Algorithmus | 6 |
| Dateneingabe (Schritt 1) | Algorithmus | 5 |
| Dateneingabe (Schritt 1) | Bedienung | 5 |
| Verlinkungen hinzufügen (Schritt 4) | Algorithmus | 5 |
| Referenz wählen (Schritt 5) | Bedienung | 5 |
| Zitat hinzufügen (Schritt 6) | Bedienung | 5 |

Umfrage Teil eins (Allgemeine Angaben zum Prototypen)

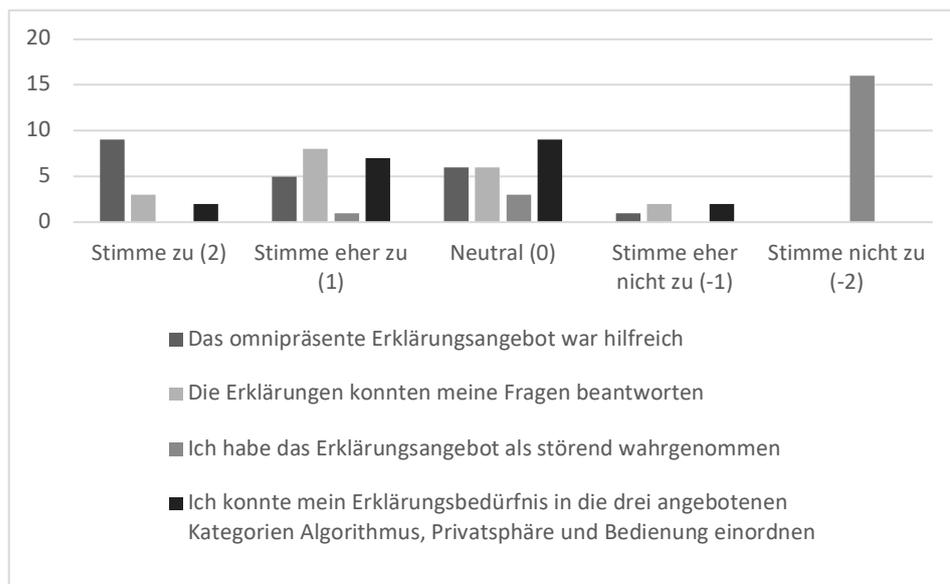


Abbildung 4.1: Ausschnitt Umfrageergebnisse Teil eins

67% der Teilnehmenden haben das Erklärungsangebot mindestens als eher hilfreich wahrgenommen (siehe Abb. 4.1). Eher hilfreich bzw. 1 entspricht dabei auch dem Mittelwert über die Aussagen aller Teilnehmenden. Lediglich eine Person fand es eher weniger hilfreich. Bei den 25% der Teilnehmenden, die sich die meisten Erklärungen haben anzeigen lassen, lag der Schnitt bei 1,8, sie haben das Angebot also deutlich als hilfreich wahrgenommen.

Auch die Platzierung des Erklärungsangebotes an der immer gleichen Stelle wurde durchschnittlich mit 1,6, also eher hilfreich bis hilfreich,

bewertet. Der Aussage, 'Ich habe das Erklärungsangebot als störend wahrgenommen' stimmten 76% der Teilnehmenden nicht und eine Person eher zu. Diese Person nahm das Erklärungsangebot auch als Einzige als eher nicht hilfreich wahr.

Die Teilnehmenden, die sich Erklärungen haben anzeigen lassen, gaben mit einer Zustimmung von 0,6 an, dass ihre Fragen mithilfe der Erklärungen eher beantwortet wurden. Die 25% der Teilnehmenden, die besonders viele Erklärungen angefordert haben, stimmten mit einem Durchschnittswert von 0,7 zu (vgl. Abb. 4.1).

Zehn Prozent der Teilnehmenden gaben an, dass sie ihre Fragen eher nicht in die drei angebotenen Kategorien einordnen konnten. Eine Person dieser Gruppe gab explizit an, die Kategorie *Algorithmus* nicht verstanden, bzw. sie nicht als notwendig empfunden zu haben.

81% der Teilnehmenden gaben mindestens an, sich ein solches Erklärungsangebot auch in anderer Software eher zu wünschen. Die durchschnittliche Zustimmung zu diesem Wunsch lag bei einem Wert von 1 und bei den 25% der Teilnehmenden, die sich besonders viele Erklärungen haben anzeigen lassen, bei 1,5. Die Standardabweichung war bei dieser Frage mit einem Wert von 0,75 besonders niedrig, die Teilnehmenden waren sich hier also überdurchschnittlich einig.

Umfrage Teil zwei (Angaben zu den Systemaspekten)

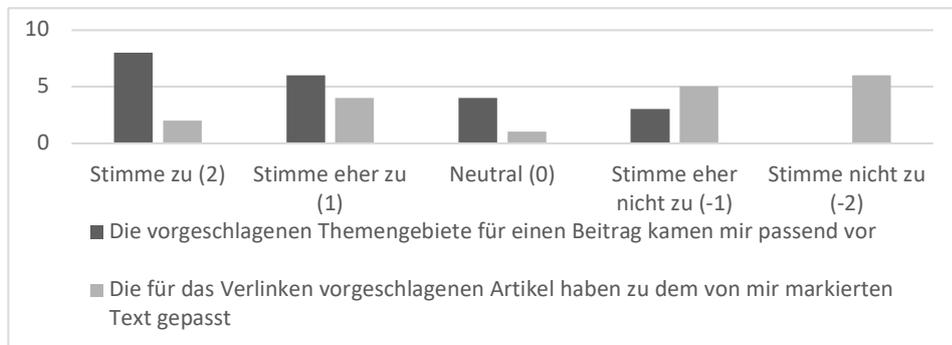


Abbildung 4.2: Ausschnitt Umfrageergebnisse Teil zwei (Algorithmus)

Die Zustimmung zu der Frage, ob die Vorschläge der Software als passend empfunden wurden, lag bei den Themenvorschlägen bei 1 und bei den vorgeschlagenen Referenzen bei -0,5 (vgl. Abb. 4.2). In beiden Fällen wurden zwei Erklärungen zum Algorithmus angezeigt. Im ersten Fall von zwei Teilnehmenden, die die Themenvorschläge als eher passend und im zweiten Fall von zwei Teilnehmenden, die die Referenzen als nicht passend wahrgenommen haben. Dabei gaben sie an, dass die Erklärungen ihnen dabei geholfen haben, die Software eher besser zu verstehen.

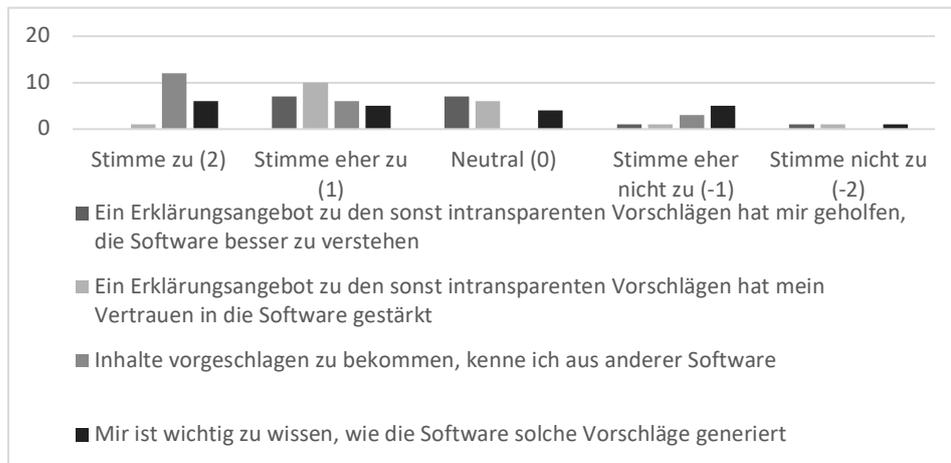


Abbildung 4.3: Ausschnitt Umfrageergebnisse Teil zwei (Algorithmus)

Unabhängig davon, ob Erklärungen angefordert wurden, gab es einen durchschnittlichen Zustimmungswert von 0,5 zu der Aussage, 'Ein Erklärungsangebot zu den sonst intransparenten Vorschlägen hat mein Vertrauen in die Software gestärkt' (vgl. Abb. 4.3). 52% der Teilnehmenden haben der Aussage dabei zugestimmt.

Mit einem Zustimmungswert von 1,3 gaben die Teilnehmenden an, bereits aus anderer Software eher zu kennen, dass Inhalte vorgeschlagen werden. Sämtliche der 14% der Teilnehmenden, die mit solchen Vorschlägen eher noch nicht konfrontiert wurden, fragten Erklärungen aus dem Bereich Algorithmus an. Bei den Teilnehmenden, die Vorschläge eher gewohnt sind, waren es 50%. Das Interesse daran, wie diese Vorschläge generiert werden, fällt mit einer Zustimmung von 0,47 eher neutral aus. Die Teilnehmenden mit Interesse an dieser Information ließen sich durchschnittlich ebenfalls 1,7, die anderen 0,6 Erklärungen aus dem Bereich Algorithmus anzeigen.

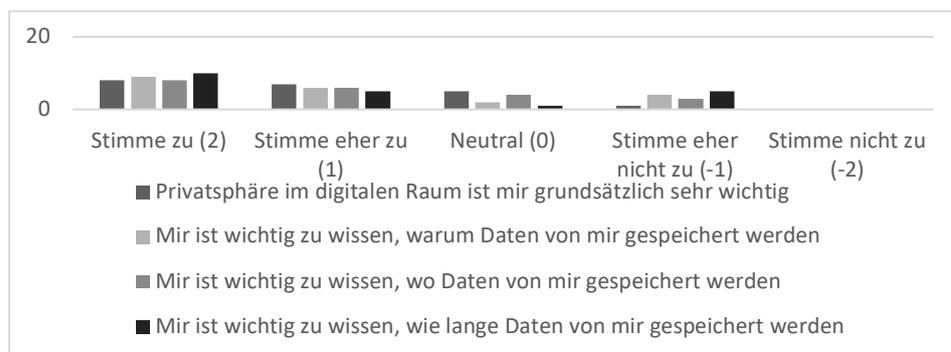


Abbildung 4.4: Ausschnitt Umfrageergebnisse Teil zwei (Privatsphäre)

Zum Thema Privatsphäre geben die Teilnehmenden durchschnittlich an, dass ihnen Privatsphäre im digitalen Raum eher wichtig (Schnitt 1), 38% sogar dass sie ihnen sehr wichtig ist (vgl. Abb.4.4). 29% standen dieser Aussage neutral oder eher ablehnend gegenüber. Ähnlich war die Zustimmung dazu, dass es wichtig ist, warum, wo und wie lange Daten gespeichert werden. Bei diesen drei Fragen lag die Zustimmung bei 0,9-1, allerdings bei einer bis zu 0,3 höheren Standardabweichung von 1,1-1,2.

Teilnehmende, die maximal angaben, dass ihnen Privatsphäre im digitalen Raum weniger wichtig ist, forderten keine Erklärungen, Teilnehmende die mindestens eher wichtig angegeben haben, durchschnittlich 1,3 Erklärungen zu diesem Systemaspekt an. Besonders vertreten waren hier auch die Teilnehmenden, die nur 'neutral' angegeben hatten. Diese rund 24% ließen sich durchschnittlich 2,2 Erklärungen zu diesem Systemaspekt anzeigen.

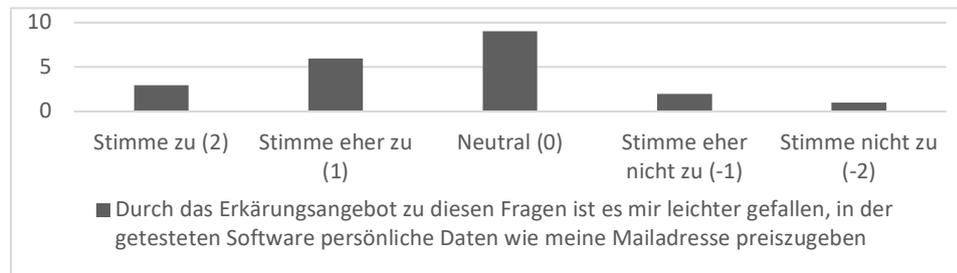


Abbildung 4.5: Ausschnitt Umfrageergebnisse Teil zwei (Privatsphäre)

Die Zustimmung zu der Aussage, dass Erklärungen zu dem Systemaspekt *Privatsphäre* das Vertrauen erhöht haben, lag bei 0,4 (vgl. Abb.4.5). Bei den 25% der Nutzenden, die sich besonders viele Erklärungen zu diesem Systemaspekt anzeigen ließen, lag die Zustimmung bei 0,8. Unabhängig von der Erklärungsanfrage lag die Zustimmung zu der Aussage bei den Teilnehmenden, denen Privatsphäre im digitalen Raum besonders wichtig ist, bei 1.

Bei 13 bzw. 62% der Teilnehmenden sind im Laufe des Prozesses Probleme aufgetreten. Diese Probleme traten an folgenden Stellen auf:

- Nach dem Auswählen einer Referenz aus der Liste mit Vorschlägen in (Schritt 5) wurde zwar das Textfeld mit dem Link gefüllt, das Feld und der 'Weiter'-Button waren aber aus dem Sichtfeld gescrollt, sodass dem Teilnehmenden nicht klar war, dass der 'Weiter'-Button nun aktiviert ist.

5 Mal aufgetreten. In vier davon wurde eine Erklärung angefragt.

- Starten mit der Aufgabenstellung war nicht möglich, weil nicht alle Erklärungskategorien geöffnet wurden.

4 Mal aufgetreten. Die Erklärung hierfür stand bereits im Text.

- Dem Teilnehmenden war unklar, dass eine Referenz durch Auswählen von Text hinzugefügt werden kann.

3 Mal aufgetreten. In allen Fällen wurde eine Erklärung angefordert, wobei eine Person die Fragestellung fälschlicher Weise dem Systemaspekt *Algorithmus* statt *Bedienung* zugeordnet hat.

- Die eingegebene Mailadresse wurde von der Software nicht akzeptiert. (Enthielt nicht erlaubte Sonderzeichen)

Ein Mal aufgetreten, wurde über Ausprobieren gelöst.

62% der Teilnehmenden geben an, dass sie das Starten von Aktionen über das Auswählen von Text nicht oder eher nicht kannten. 53% dieser Personen fragten bei der entsprechenden Ansicht eine Erklärung an. Damit kommen 70% der Erklärungsanfragen auf dieser Ansicht von dieser Personengruppe.

29% der Teilnehmenden nahmen die Software im Ganzen als eher unintuitiv wahr. Wider Erwarten fragte diese Personengruppe weniger Erklärungen an als der Gesamtdurchschnitt der Teilnehmenden. Die Hälfte dieser Personengruppe stufte sich selbst als technisch weniger versiert ein.

Unabhängig von den oben thematisierten Fehlern gaben 57% der Teilnehmenden an, dass an mindestens einer Stelle im Prozess der 'Weiter'-Button aus für sie unverständlichen Gründen deaktiviert war. 50% davon nahmen die Erklärung als hilfreich war.

Kapitel 5

Diskussion

Mit den im letzten Kapitel ausgearbeiteten Ergebnissen der Nutzerstudie werden im Folgenden die Forschungsfragen beantwortet.

5.1 Beantwortung von RQ1

RQ1 Wie groß ist der tatsächliche Erklärungsbedarf des Nutzens in Bezug auf die Systemaspekte Algorithmus, Privatsphäre und Bedienung in der hier verwendeten Beispielsoftware?

Wir nehmen an, dass die explizite Anfrage von Erklärungen Resultat eines realen Erklärungsbedarfs ist. Darauf aufbauend zeigen die Ergebnisse der Nutzerstudie bei 76% der Teilnehmenden einen Erklärungsbedarf.

Bei Personen im Alter von über 25 Jahren war dieser Erklärungsbedarf besonders hoch ausgeprägt. Das kann darauf zurückzuführen sein, dass die jüngere Generation es als *digital natives* gewohnt ist, mit vielen Informationen konfrontiert zu werden. Dabei hat sie womöglich gelernt, schnell Entscheidungen zu treffen, ohne die Informationslast durch die Anfrage neuer Informationen weiter zu erhöhen. Die ältere Generation könnte stattdessen planvollere Entscheidungen anstreben und dafür eine größere Informationsbasis wünschen.

Ebenfalls höher ist der Erklärungsbedarf von technisch weniger versierten Personen gegenüber technisch versierten Personen. Dabei zeichnet sich aber ab, dass technisch versiertere Personen, die Erklärungen anfragen, dann durchschnittlich häufiger Erklärungen fordern. Das lässt auf eine konsequentere Softwarenutzung schließen und weist eventuell auf ein höheres generelles Interesse an der genutzten Software hin, während unsichere Softwarenutzende Unterstützungsangebote eher in für sie problematischen Anwendungsfällen ausprobieren. Der Unterschied in der Anzahl angezeigter Erklärungen ist mit Blick auf die Datenmenge allerdings nicht statistisch signifikant.

Besonders hoch ist der Erklärungsbedarf der Nutzerstudie zufolge im Bereich der Bedienung. Alle Teilnehmenden, die sich Erklärungen anzeigen ließen, forderten auch mindestens eine Erklärung aus dem Bereich Bedienung. Auch im Bereich Algorithmus und Privatsphäre haben sich aber über die Hälfte der Nutzenden Erklärungen anzeigen lassen. Sowohl die Anzahl an Personen, die sich hier Erklärungen anzeigen lassen haben, als auch die Menge der angezeigten Erklärungen waren in diesen Bereichen aber kleiner.

Bei der Mehrheit der Teilnehmenden zeichnet sich also ein Erklärungsbedarf ab. Dieser ist bei tendenziell älteren oder technisch weniger versierten Personen höher ausgeprägt. In höherem Maße von dem Erklärungsbedarf betroffen ist der Systemaspekt Bedienung.

5.2 Beantwortung von RQ2

RQ2 In welchen Teilszenarien sind Erklärungen zu welchen Systemaspekten relevant?

Bedienung

Besonders viele Erklärungen wurden auf der Ansicht 'Verlinkungen hinzufügen' im Bereich des Softwareaspekts Bedienung angezeigt. Knapp die Hälfte der Teilnehmenden öffneten diese Erklärungen. Auf dieser Ansicht traten am dritthäufigsten Probleme auf. Diese Ansicht zeichnete sich durch eine neuartige Bedienung aus. Der Teilnehmende sollte eine Aktion starten, indem er Text markiert. 62% der Teilnehmenden bestätigten, dass ihnen diese Interaktionsform unbekannt war. Das zeigt, dass Erklärungen zu diesem Systemaspekt besonders dann gefragt sind, wenn Nutzende mit einem für sie neuartigen Bedienungskonzept konfrontiert werden, für das sie noch kein mentales Modell entwickelt haben.

Knapp 80% der Teilnehmenden, die beim Auswählen einer Referenz für einen markierten Text Probleme hatten, ließen sich dazu eine Erklärung anzeigen. Beim Eingeben der E-Mailadresse äußerte nur eine Person Probleme. Diese Person löste das Problem durch Ausprobieren. Neben neuartigen Bedienungskonzepten und inhaltlichen Erklärungen zu den Eingabemöglichkeiten könnten Erklärungen für Nutzende demzufolge auch dann interessant sein, wenn Probleme auftreten, die sich nicht durch Ausprobieren lösen lassen.

Zu dem gleichen Systemaspekt ließen sich acht Nutzende in der Ansicht 'Beitrag schreiben' eine Erklärung anzeigen. Die Aufgabenstellung stellte keine Anforderungen an den Text, abgesehen von Thema und Länge gab es keine Vorgaben. Vermutlich war den Teilnehmenden hier nicht klar, wie der zu schreibende Text aussehen, ob es sich beispielsweise um einen wissenschaftlichen Text handeln soll. Unter dem Systemaspekt Bedienung erwarten Nutzende demnach auch dann Erklärungen, wenn für sie unklar

ist, wie Eingaben inhaltlich aussehen sollen. Möglicherweise ist dafür die Einführung einer neuen Erklärungskategorie sinnvoll.

5.2.1 Privatsphäre

Zur Untersuchung des Erklärungsbedarfs von Privatsphäre wurde primär die Dateneingabe im ersten Schritt des Prototypen entwickelt. Mit 38% der Teilnehmenden ließen sich hier am zweithäufigsten Teilnehmende eine Erklärung anzeigen. Erklärungen zu diesem Systemaspekt sind für Teilnehmende folglich vor allem dann interessant, wenn explizit persönliche Daten von ihnen abgefragt werden. Auch bei Teilnehmenden, die nur eine neutrale Haltung zur Wichtigkeit von Privatsphäre im digitalen Raum hatten, waren diese Erklärungen gefragt. Das kann darauf hinweisen, dass diese Teilnehmenden ihr Interesse selbst unterschätzt haben. Möglicherweise haben Teilnehmende mit höherem Interesse aber auch softwareunabhängige Lösungen entwickelt, ihre Daten zu schützen, und daher nur noch ein vermindertes Interesse an Erklärungen zu diesem Aspekt. Zum Beispiel können sie die Bereitstellung persönlicher Daten reduzieren, indem temporäre E-Mail Postfächer genutzt werden, die sie eigens für solche Dienste anlegen.

Etwa halb so viele Erklärungsanfragen wurden beim Schreiben des Textes und Auswählen von Textstellen zum Hinzufügen von Referenzen gestellt. Die durch den Teilnehmenden generierten Daten wurden dementsprechend auch als persönliche Daten wahrgenommen und lösten daraufhin ein Erklärungsbedürfnis zu diesem Erklärungsaspekt aus.

5.2.2 Algorithmus

Zu diesem Systemaspekt wurde angenommen, dass generierte Vorschläge ein Erklärungsbedürfnis hervorrufen. Zur Simulation dieses Effekts wurden zwei Ansichten entworfen, die über Vorschläge verfügen. Die auf den beiden Seiten jeweils angezeigten Vorschläge wurden von den Teilnehmenden als unterschiedlich passend wahrgenommen. Auf beiden Ansichten wurden nur jeweils zwei Erklärungen zu diesem Systemaspekt angezeigt. Weder unpassende noch passende Vorschläge werfen also einen Erklärungsbedarf auf. Die Umfrageergebnisse zeigen, dass 86% der Teilnehmenden bereits häufiger mit Vorschlägen konfrontiert wurden. Das Interesse daran, wie diese Vorschläge generiert werden, wurde als neutral angegeben. Zwei Teilnehmende gaben explizit an, dass sie Erklärungen zu diesem Systemaspekt für weniger relevant halten: „Ich glaube für die Nutzer*innen von den meisten Apps ist die Kategorie Algorithmus nicht relevant.“

Möglich ist aber, dass Erklärungen zu diesem Systemaspekt als relevanter wahrgenommen werden, wenn die Vorschläge nicht nur eine Ergänzung, sondern kritischer Bestandteil des Systems sind, der Algorithmus also den Programmablauf bestimmt, oder mehr Daten mit einbezogen werden.

Die Ansichten, auf denen besonders viele Erklärungen zu dem Systemaspekt *Algorithmus* angezeigt wurden, hatten unter diesem Aspekt zumeist tatsächlich eine sehr geringe Relevanz. Auf den Ansichten 'Beitrag schreiben' oder 'Dateneingabe' zum Beispiel werden weder komplexe Berechnungen durchgeführt noch anderen Inhalte angezeigt, bei denen unklar sein könnte, wie die Software sie generiert hat.

Möglicherweise wurden die Hintergründe zu dieser Kategorie teilweise missverstanden und deshalb auch auf unpassenden Seiten Erklärungen aufgerufen. Darauf lässt der Kommentar eines Teilnehmers schließen, der anmerkte, dass er sich bei einem auftretenden Problem in der Bedienung zwar eine Erklärung hat anzeigen lassen, diese aber nicht weitergeholfen hat. Auffallend war an dieser Stelle, dass der Teilnehmer tatsächlich nicht die passenden Erklärungen zu dem Systemaspekt *Bedienung* stattdessen aber die zum *Algorithmus* angefragt hat. Das Missverständnis kann mit unzureichender Erklärung zu diesem Systemaspekt selbst zusammenhängen. Der Kommentar einer Teilnehmerin „Ich habe die Kategorie Algorithmus nicht ganz verstanden bzw. habe ihn nicht unbedingt als notwendig empfunden.“ lässt aber auch den Schluss zu, dass Teilnehmende davon ausgegangen sind, den Aspekt nicht richtig zu verstehen, weil sie keine Notwendigkeit dafür sahen. Infolge dessen kann eine falsche Annahme zu dem Hintergrund des Systemaspekts auch auf einer selbst erschlossenen Erklärung basieren, die aus der Perspektive des Teilnehmenden mehr Sinn ergab.

Die Einblendung von Vorschlägen hat im Rahmen dieses Prototypen nicht zu einem erhöhten Erklärungsbedürfnis zum Systemaspekt *Algorithmus* geführt. Unter Umständen hängt das aber mit Limitationen der Studie (Siehe 5.5) oder Unverständnis des Systemaspekts zusammen. Die Bedürfnisse, die stattdessen zu einer Forderung von Erklärungen geführt haben, lassen sich eher dem Systemaspekt *Bedienung* oder einer neu einzuführenden inhaltlichen Kategorie zuordnen.

5.3 Beantwortung von RQ3

RQ3 Lässt sich der Erklärungsbedarf mithilfe eines omnipräsenten Erklärungsangebots decken, über das auf Nachfrage Erklärungen angezeigt werden?

Die Aussage 'Die Erklärungen konnten meine Fragen beantworten' bekam nur eine verhaltene bis neutrale Zustimmung. Kommentare von den knapp 10% der Teilnehmenden, die eher nicht zustimmten, deuten dabei auf inhaltliche Fragen hin. Auf Fragen zur Konzeption der Software, beispielsweise dem Sinn von Zitaten als Ergänzung zu den Referenzen, war der Erklärer nicht ausgelegt. Das am zweithäufigsten aufgetretene Problem war, dass Teilnehmende den Text auf der Startseite nicht ausreichend lasen

und deshalb nicht wussten, das zum Starten erst die drei Erklärungskategorien geöffnet werden müssen. Für einige Teilnehmende war die Menge an Informationen, die zur Einführung in das Softwarekonzept bereitgestellt wurden, demnach bereits in der verwendeten Version zu groß. Weitere Ausführungen an dieser Stelle können insofern nicht die Lösung sein. Eine Einbettung eines Erklärungsaspekts, der auf Anforderung des Nutzenden hin inhaltliche Fragen zu dem aktuell durchzuführenden Schritt beantwortet, kann in zukünftigen Arbeiten also ein guter Schritt sein.

Unter den Systemaspekten *Privatsphäre* und *Bedienung* wurden Erklärungen als hilfreich wahrgenommen. Daraus folgt, dass das Erklärungsmodell an sich dazu geeignet ist, auftretende Fragestellungen zu beantworten, solange sich diese Fragestellungen in dem verwendeten Systemumfeld in die für die Erklärungskomponente gewählten Kategorien einordnen lassen.

Nur 30% derer, die auf der, dem Erklärungsbedarf zufolge problematischsten Ansicht eine Erklärung anforderten und folglich Schwierigkeiten gehabt zu haben scheinen, gaben in der Umfrage an, dass dort ein Problem aufgetreten ist. Das deutet darauf hin, dass mithilfe von Erklärungen gelöste Bedienungsprobleme nur selten nachhaltig als Problem angesehen werden.

Eine Teilnehmerin erwähnte in einem Kommentar die Vorschläge von „ChatGPT“, traf also eine falsche, aber für sie ausreichende Annahme über den Ursprung der Vorschläge, ohne dafür Erklärungen anzufordern. Dass Teilnehmende sich Fragestellungen selbst erklären können, ohne dafür weitere Informationen aufnehmen und verarbeiten zu müssen, ist ein Vorteil des latenten Erklärungsangebotes. Eigenes Erschließen spart oft Aufwand und in dem Moment, wo sich trotzdem Fragen stellen, stehen die Erklärungen zur Verfügung. Gleichzeitig können dabei, wie in diesem Beispiel, falsche Annahmen entstehen. Was in diesem Fall harmlos ist, kann in anderen Szenarien das Vertrauen in Software korrumpieren.

Sollte eine Erklärung zu einem Teilszenario unter einem für das Vertrauen in Software besonders relevanten Systemaspekt wie *Privatsphäre* oder *Algorithmus* besonders entscheidend sein, sollten diese Erklärungen daher nicht erst auf Anfrage hin angezeigt werden. Das gilt insbesondere dann, wenn die Erklärung entscheidend von dem abweicht, was sich der durchschnittliche Nutzende in dem entsprechenden Anwendungsfall selbst erschließt.

5.4 Beantwortung von RQ4

RQ4 Werden latente Erklärungen, die erst auf Abfrage bereitgestellt werden, in der Benutzung von Software als hilfreich und vertrauensaufbauend wahrgenommen?

Die Mehrzahl der Teilnehmenden hat das Erklärungsangebot als hilfreich wahrgenommen. Das Vertrauen in die Software konnte durch das Erklärungsangebot in Bezug auf Privatsphäre insbesondere bei den Personen,

die besonderen Wert auf die Privatsphäre im digitalen Raum legen, erhöht werden. Die Teilnehmenden, die sich unter diesem Systemaspekt besonders viele Erklärungen anzeigen ließen, stimmten hier in geringerem Maße durchschnittlich ebenfalls zu. Statt der Menge an angeforderter Erklärungen ist folglich die Wahrnehmung der Notwendigkeit von Privatsphäre und das reine Angebot an Erklärungen der ausschlaggebende Faktor.

Insgesamt wurde der vertrauens erhöhenden Effekt von latenten Erklärungen mit neutral bis eher positiv bewertet. Da das Erklärungsangebot nur von einer Person als eher störend empfunden wurde, ist es weiterhin als Mehrwert anzunehmen. Dass eine große Mehrheit der Teilnehmenden sich dieses auch in anderer Software wünscht, unterstützt diese Annahme.

5.5 Limitationen der Studie

Einige Einflussfaktoren können die Aussagekraft der Ergebnisse dieser Arbeit möglicherweise einschränken. Dabei sind insbesondere die Teilnehmenden, der entwickelte Prototyp und die angeschlossene Umfrage zu betrachten.

Die Studie hatte 21 Teilnehmende. Diese geringe Teilnehmerzahl kann bedeuten, dass die Ergebnisse nicht repräsentativ sind. Auch die demographischen Daten der Teilnehmenden können die Repräsentativität der Ergebnisse einschränken. Hier war die Altersgruppe der über 30 Jährigen mit 14% eventuell unter- und der Anteil an Personen mit technischem beruflichen Hintergrund mit 48% möglicherweise überrepräsentiert.

In dem für den Prototypen gewählten Szenario könnten für die einzelnen Systemaspekte entwickelte Teilszenarien ungleich ausgeprägt sein. Unter dem Systemaspekt *Privatsphäre* ist es schwierig, übermäßige Datensammlung zu simulieren, da keine so lange und persönliche Nutzung erfolgt wie beispielsweise bei der Nutzung von sozialen Medien. Auch unter dem Aspekt *Algorithmus* ist der Effekt der intransparenten Vorschläge, mit denen Nutzende dort häufig konfrontiert sind, schwierig abzubilden. Unter diesem Aspekt kommt hinzu, dass Ergebnisse des Prototypen höchstens die weitere Nutzung beeinflussen, darüber hinaus aber keinen Effekt haben. Außerdem ist eine Abweichung einer Ausgabe von der Gewohnten schwierig abzubilden, ohne diese durch eine Langzeitnutzung etabliert zu haben. Hier ist der Erklärungsbedarf, der beispielsweise bei einem ungewöhnlichen Routen-Vorschlag einer Navigationssoftware auftreten kann, unter Umständen nicht abgedeckt. Bei der Konzeption wurde allerdings darauf geachtet, auf möglichst vielen Ansichten Inhalte zu platzieren, zu denen zu möglichst allen Systemaspekten Erklärungen Sinn ergeben, um eine Ausgewogenheit zu erreichen.

Die in der Umfrage gestellten Fragen wurden von dem Studienautor formuliert, von dessen Betreuerin überprüft und in einer Pilotstudie an zwei Teilnehmenden getestet. Eine Objektivität der Fragen kann aber nicht garantiert werden.

Kapitel 6

Zusammenfassung und Ausblick

In diesem Kapitel wird ein abschließender Überblick über die Inhalte und die wichtigsten Ergebnisse dieser Arbeit gegeben.

6.1 Zusammenfassung

Um den Erklärungsbedarf von Software untersuchen zu können, wurde in dieser Arbeit ein Softwarekonzept entwickelt und prototypisch umgesetzt. Der Prototyp dieses Konzepts implementiert einen omnipräsenten Erklärer, über den bei Bedarf Erklärungen angezeigt werden können. Das Erklärungen erst auf Anfrage angezeigt werden, bietet auf der einen Seite die Möglichkeit, den Erklärungsbedarf zu messen, und schränkt den Nutzenden auf der anderen Seite nicht durch die Anzeige von zu vielen kognitiv belastenden Informationen in der Bedienung ein. Die Erklärungen sind, basierend auf Erkenntnissen in der Literatur, in die drei Systemaspekte *Algorithmus*, *Privatsphäre* und *Bedienung* untergliedert.

In einer Nutzerstudie sollten Teilnehmende eine Aufgabenstellung in diesem Prototypen lösen. Im Anschluss wurden in einer Umfrage deren Reaktionen auf die Erklärungskomponente ermittelt. Um authentische Verhaltensweisen beobachten zu können, wird der tatsächliche Hintergrund der Studie nicht erklärt und dem Teilnehmenden stattdessen suggeriert, dass Inhalte des eigens entwickelten Softwarekonzepts untersucht werden, das neue Ansätze im Bereich der digitalen Kommunikation verfolgt.

Dabei fragten über die Hälfte der Teilnehmenden Erklärungen an. Daraus kann ein bei über der Hälfte der Teilnehmenden aufgetretener Erklärungsbedarf geschlossen werden. Eine Komponente, die erst auf Anfrage Erklärungen anzeigte, erwies sich als geeignetes Mittel, um diesen Bedarf zu decken, ohne überflüssige kognitiv belastenden Informationen anzuzeigen, welche Nutzende in ihrer Softwarenutzung beeinträchtigen.

Ein besonders hohes Erklärungsbedürfnis äußerten die Teilnehmenden unter dem Systemaspekt *Bedienung*, primär auf Ansichten, die neuartige Bedienungsverfahren verwendeten. Mit Erklärungen gelöste Bedienungsschwierigkeiten wurden in nur 30% der Fälle als Probleme wahrgenommen. Auch Erklärungen zu dem Aspekt *Privatsphäre* wurden vermehrt angefragt, wenn persönliche Daten von Teilnehmenden bereitzustellen waren. Selbstverfasste Inhalte wurden dabei ebenfalls als persönliche Daten wahrgenommen. Der Systemaspekt *Algorithmus* hingegen wurde wenig genutzt, mit dem Aspekt *Bedienung* verwechselt oder angefragt, wenn sich inhaltliche Fragen zu dem gerade durchzuführenden Schritt stellten. In diesen Aspekt einzuordnende Informationen scheinen einige Nutzende nicht zu interessieren, andere erschließen sich diese selbst.

Schon das alleinige Angebot von Erklärungen wurde den Umfrageergebnissen zufolge als positiv wahrgenommen und konnte in einigen Fällen zum Aufbau von Vertrauen beitragen.

6.2 Ausblick

Die Ergebnisse zeigen, dass die Systemaspekte, nach denen Erklärungen in dieser Arbeit kategorisiert wurden, noch nicht optimal gewählt sind. Stattdessen deuten sie darauf hin, dass der Systemaspekt *Algorithmus* in den Aspekt *Bedienung* integriert werden könnte und an dessen Stelle ein Aspekt, der inhaltliche Fragen zum aktuell durchzuführenden Schritt behandelt, Sinn ergibt. An der Auswahl sollte in zukünftigen Arbeiten folglich weiter geforscht werden.

Zu bedenken ist dabei, dass differente Software unter Umständen Erklärungsbedarf zu differenten Systemaspekten auslöst. Dementsprechend sollte in zukünftigen Arbeiten erforscht werden, ob es Szenarien gibt, in denen Nutzende Informationen zu dem Systemaspekt *Algorithmus* mehr Interesse entgegenbringen und welchen Effekt latente Erklärungen dann zeigen. Das können Szenarien sein, in denen die Ergebnisse des Algorithmusses weitreichendere Konsequenzen haben oder ungewöhnlicher wirken, als das in dieser Arbeit der Fall war.

Auch die weitgehend neutrale Bewertung des Effekts der Erklärungskomponente auf das Vertrauen in die Software kann darauf zurückzuführen sein, dass das gewählte Szenario unter den Systemaspekten *Privatsphäre* und *Algorithmus* für den Nutzenden wenig herausfordernd war. Wenn die Ergebnisse des Algorithmusses weitreichendere Konsequenzen haben und persönliche Daten intransparenter verarbeitet werden, könnten hier deutlichere Ergebnisse auftreten.

Den Studienergebnissen zufolge können Nutzende Falschannahmen über die Eigenschaften von Software treffen, wenn korrigierende Informationen nur auf Nachfrage bereitgestellt werden. Interessant ist daher weiterhin, ob

diese Falschannahmen Vertrauen in Software korrumpieren können, oder ob Nutzende diese durch Anfrage von Erklärungen validieren, bevor sie Konsequenzen daraus ziehen.

Die Ergebnisse der Nutzerstudie sollten aufgrund der geringen Teilnehmerzahl zudem für zukünftige Arbeiten als Annahmen genutzt und dort verifiziert werden. Insbesondere die Einflüsse auf den Erklärungsbedarf durch das Alter und die technische Versiertheit der Teilnehmenden kann mit höheren Teilnehmerzahlen sicherer aufgeschlüsselt werden.

Anhang A

Ein Anhang

A.1 Umfrage

Die folgenden Fragen wurden in der Umfrage nach der Bearbeitung der Aufgabenstellung in dem Prototypen gestellt. Einige der Fragen wurden in Abhängigkeit von der Antwort auf andere Fragen angezeigt.

Fragen zu dem getesteten Prototypen

Mit diesen Fragen möchten wir herausfinden, wie Sie das Erklärungsangebot wahrgenommen haben.

- Haben Sie das Erklärungsangebot in Anspruch genommen, sich also Erklärungen einblenden lassen? [Ja, Nein]
- Warum haben Sie sich keine Erklärungen anzeigen lassen?
 - [Ich hatte keine Fragen, für die ich eine Erklärung brauchte]
 - [Ich konnte die Fragen, die ich hatte nicht in die angebotenen Kategorien einordnen]
 - [Das Erklärungsangebot war nicht präsent genug]
 - [Sonstiges – Freitext]
- Inwieweit stimmen Sie folgenden Aussagen über das Erklärungsangebot der getesteten Software zu?
 - jeweils [Stimme zu, Stimme eher zu, neutral, Stimme eher nicht zu, stimme nicht zu]
 - Das omnipräsente Erklärungsangebot (Fragezeichen unten Rechts in der Ecke der Software) war hilfreich
 - Es ist hilfreich, immer an der gleichen Stelle Hilfestellungen finden zu können

- Die Erklärungen konnten meine Fragen beantworten
- Ich habe das Erklärungsangebot als störend wahrgenommen
- Ich konnte mein Erklärungsbedürfnis in die drei angebotenen Kategorien Algorithmus, Privatsphäre und Bedienung einordnen
- Ich wünsche mir einen solchen Erklärer auch in anderer Software
- Welche Fragestellung blieb unbeantwortet? [Freitext]
- Was hat sie an dem Erklärungsangebot gestört? [Freitext]
- Welche Fragestellung ließ sich nicht in die gegebenen Kategorien einordnen? Haben Sie einen Vorschlag, für eine passende Kategorie? [Freitext]
- Weitere Anmerkungen [Freitext]

Kategorien des Erklärungsangebots

In diesem zweiten Schritt möchten wir herausfinden, unter welchen Voraussetzungen Nutzende bevorzugt Erklärungen einblenden oder darauf verzichten.

- Inwieweit stimmen Sie folgenden Aussagen über Ihre Wahrnehmung der Ergebnisse der getesteten Software zu?
jeweils [Stimme zu, Stimme eher zu, neutral, Stimme eher nicht zu, stimme nicht zu]
 - Die vorgeschlagenen Themengebiete für einen Beitrag kamen mir passend vor
 - Die für das Verlinken vorgeschlagenen Artikel haben zu dem von mir markierten Text gepasst
 - Ein Erklärungsangebot zu den sonst intransparenten Vorschlägen hat mir geholfen, die Software besser zu verstehen
 - Ein Erklärungsangebot zu den sonst intransparenten Vorschlägen hat mein Vertrauen in die Software gestärkt (Unabhängig davon, ob das Angebot auch tatsächlich in Anspruch genommen wurde)
 - Inhalte vorgeschlagen zu bekommen, kenne ich aus anderer Software (Videos, die mir gefallen könnten; Personen, denen ich vielleicht folgen möchte; ...)
 - Mir ist wichtig zu wissen, wie die Software solche Vorschläge generiert
- Inwieweit stimmen Sie folgenden Aussagen über Privatsphäre zu?
jeweils [Stimme zu, Stimme eher zu, neutral, Stimme eher nicht zu, stimme nicht zu]

- Privatsphäre im digitalen Raum ist mir grundsätzlich sehr wichtig
- Mir ist wichtig zu wissen, warum Daten von mir gespeichert werden
- Mir ist wichtig zu wissen, wo Daten von mir gespeichert werden
- Mir ist wichtig zu wissen, wie lange Daten von mir gespeichert werden
- Durch das Erklärungsangebot zu diesen Fragen ist es mir leichter gefallen, in der getesteten Software persönliche Daten wie meine Mailadresse preiszugeben (Unabhängig davon, ob das Angebot tatsächlich in Anspruch genommen wurde)
- Inwieweit stimmen Sie folgenden Aussagen über die Bedienung der getesteten Software zu?
 - jeweils [Stimme zu, Stimme eher zu, neutral, Stimme eher nicht zu, stimme nicht zu]
 - Bei der Bedienung sind Fehler aufgetreten (Fehlermeldung, rote Umrandung, ...)
 - Wenn Fehler aufgetreten sind: Das Erklärungsangebot hat mir geholfen, diese Fehler zu verstehen und zu beheben
 - Für mich war an mindestens einer Stelle unklar, warum ich den 'Weiter' Button nicht betätigen konnte
 - Wenn es Unklarheiten gab: Das Erklärungsangebot hat mir geholfen, herauszufinden, wie ich weiterkommen kann
 - Die Bedienung der Software war intuitiv für mich
 - Eine Aktion zu starten, indem ich Text markiere, konnte ich aus anderer Software
- An welcher Stelle sind Fehler aufgetreten oder war unklar, warum es nicht weitergeht? Wie haben Sie das Problem gelöst? [Freitext]
- Weitere Anmerkungen [Freitext]

Demographische Fragen

- Wie alt sind Sie? [Zahl]
- Welchem Geschlecht ordnen sie sich zu? [männlich, weiblich, divers]
- Welcher Tätigkeit gehen sie gerade nach? [Ausbildung / Studium, Arbeit; jeweils mit Kommentarmöglichkeit für Tätigkeitsfeld]

- Inwieweit stimmen Sie folgenden Aussagen über Ihre Softwarenutzung zu?
jeweils [Stimme zu, Stimme eher zu, neutral, Stimme eher nicht zu, stimme nicht zu]
 - Ich arbeite im beruflichen und oder privaten Kontext viel mit verschiedener Software
 - Ich fühle mich grundsätzlich sicher in der Bedienung von Software

Literaturverzeichnis

- [1] A. Adadi and M. Berrada. Peeking inside the black-box: a survey on explainable artificial intelligence (xai). *IEEE access*, 6:52138–52160, 2018.
- [2] W. Brunotte, A. Specht, L. Chazette, and K. Schneider. Privacy explanations – a means to end-user trust. *Journal of Systems and Software*, 195:111545, 2023.
- [3] L. Chazette, W. Brunotte, and T. Speith. Exploring explainability: A definition, a model, and a knowledge catalogue. In *2021 IEEE 29th International Requirements Engineering Conference (RE)*, pages 197–208, 2021.
- [4] L. Chazette and K. Schneider. Explainability as a non-functional requirement: challenges and recommendations. *Requirements Engineering*, 25(4):493–514, 2020.
- [5] R. Cummings, G. Kaptchuk, and E. M. Redmiles. ‘i need a better description’: An investigation into user expectations for differential privacy. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security, CCS ’21*, page 3037–3052, New York, NY, USA, 2021. Association for Computing Machinery.
- [6] L. M. Cysneiros, M. Raffi, and J. C. S. do Prado Leite. Software transparency as a key requirement for self-driving cars. In *2018 IEEE 26th international requirements engineering conference (RE)*, pages 382–387. IEEE, 2018.
- [7] R. Davis and B. G. Buchanan. Meta-level knowledge: Overview and applications. In *Ijcai*, volume 77, pages 920–927, 1977.
- [8] D. J. Houghton and A. N. Joinson. Privacy, social network sites, and social relations. *Journal of technology in human services*, 28(1-2):74–94, 2010.
- [9] M. A. Köhl, K. Baum, M. Langer, D. Oster, T. Speith, and D. Bohlender. Explainability as a non-functional requirement. In *2019*

- IEEE 27th International Requirements Engineering Conference (RE)*, pages 363–368. IEEE, 2019.
- [10] M. A. Köhl, K. Baum, M. Langer, D. Oster, T. Speith, and D. Bohlender. Explainability as a non-functional requirement. In *2019 IEEE 27th International Requirements Engineering Conference (RE)*, pages 363–368, 2019.
- [11] B. Lepri, N. Oliver, E. Letouzé, A. Pentland, and P. Vinck. Fair, transparent, and accountable algorithmic decision-making processes: The premise, the proposed solutions, and the open challenges. *Philosophy & Technology*, 31:611–627, 2018.
- [12] T. Lombrozo. The structure and function of explanations. *Trends in cognitive sciences*, 10(10):464–470, 2006.
- [13] A. M. McDonald and L. F. Cranor. The cost of reading privacy policies. *Isjlp*, 4:543, 2008.
- [14] T. Miller. Explanation in artificial intelligence: Insights from the social sciences. *Artificial intelligence*, 267:1–38, 2019.
- [15] A. Panesar. Machine learning and ai for healthcare: Big data for improved health outcomes. berkeley, ca: Apress; 2019. *Machine Learning and AI Ethics*, pages 207–247, 2019.
- [16] J. Pearl and D. Mackenzie. The new science of cause and effect. In *The book of why*. Basic Books New York City, 2019.
- [17] J. G. Puglisi. Developing personas for explainability requirements by means of a user study. Bachelorarbeit, Gottfried Wilhelm Leibniz Universität Hannover, FG Software Engineering, 2023.
- [18] C.-H. Tsai and P. Brusilovsky. Explaining recommendations in an interactive hybrid social recommender. In *Proceedings of the 24th International Conference on Intelligent User Interfaces, IUI '19*, page 391–396, New York, NY, USA, 2019. Association for Computing Machinery.