Leibniz Universität Hannover

Von Requirements zu Privacy Explanations: Ein nutzerzentrierter Ansatz für Usable Privacy

Der Fakultät für Elektrotechnik und Informatik

der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des akademischen Grades

Doktor rerum naturalium

(abgekürzt: Dr. rer. nat.)

vorgelegte Dissertation von Herrn

M. Sc. Wasja Brunotte

2024

1. Referent: Prof. Dr. Kurt Schneider

2. Korreferent: Prof. Dr. Andreas Vogelsang

3. Korreferent: Prof. Dr. Johannes Krugel

Vorsitzender der Prüfungskommission: Prof. Dr. Henning Wachsmuth

Tag der Promotion: 16. Februar 2024

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

Abkürzungsverzeichnis

CCPA California Consumer Privacy Act

CPRA California Privacy Rights Act

DSE Datenschutzerklärung

DSGVO Datenschutzgrundverordnung

DSR Design Science Research

EU Europäische Union

FIPP Fair Information Practice Principle

GG Grundgesetz

GT Grounded Theory

HCI Human-Computer Interaction

ISO International Standard Organization

ISR Information System Research

IT Informationstechnologie

LR Literaturrecherche

ML Maschinelles Lernen

NFR nicht-funktionale Anforderung

PbD Privacy by Design

PII Personally Identifiable Information

PSI Privacy Segmentation Index

RE Requirements Engineering

RQ Forschungsfrage

SE Software Engineering

SLR systematische Literaturrecherche

SUS System Usability Scale

UCD User-Centered Design

UI User Interface

USA United States of America

UX User Experience

XAI eXplainable Artificial Intelligence

Geduld, Hoffnung, Zuversicht und vor allem Dankbarkeit. Ich war nie allein...

...für meine Frau, Familie und Freunde – A - RIAMH NA AONAR.

Danksagungen

„Kreativität ist das, was man einsetzt, wenn man nicht weiß, was genau dabei herauskommt.“ Eine Promotion ist eine spannende und abenteuerliche Reise, auf der man, wenn man mutig und neugierig ist, viel erleben und entdecken darf. Vor allem über sich selbst. Des öfteren scheint man sich in einer Sackgasse wieder zu finden, aber auf wundersame Weise verschieben sich plötzlich Wände – und Türen öffnen sich, so dass man einen komplett neuen Blickwinkel gewinnt und aus der Sackgasse wird plötzlich eine Kreuzung vieler ungeahnter Möglichkeiten.

Demütig und dankbar bin ich für alles. Diese Reise schenkte mir meine wundervolle Frau. Eine Forscherin und Informatikerin voller Leidenschaft und Herz. Ich durfte viel von ihr lernen. Danken möchte ich auch all meinen Kolleginnen und Kollegen, Mitautorinnen und Mitautoren sowie Studierenden ohne die, diese Arbeit kaum möglich gewesen wäre. Timo Speith, für neue Perspektiven. Nils Prenner für sein wertvolles Feedback. Alexander Specht – Du hast viel zu spät begonnen. Jakob Droste für tief schürfendes miteinander Forschen. Ein besonderer Dank geht an meinen Doktorvater, Prof. Dr. Kurt Schneider, seine Loyalität, Güte und Unterstützung. Ich würde mir keinen Anderen wünschen.

Abstract

In the era of ongoing digitalization, where technology increasingly infiltrates our society, fundamental human values such as ethics, fairness, privacy, and trust have taken center stage. Digital systems have seamlessly penetrated both personal and professional spheres, offering users swift connectivity, information access, and assistance in their daily routines. In exchange, users willingly share copious amounts of personal data with these systems. However, this data collection means that that users' privacy sphere is increasingly at stake. Therefore, educating users about the information being collected and its subsequent processing is key to protect users' privacy sphere.

Legislation has established privacy policies as a means of communicating data practices. Unfortunately, these documents often prove fruitless for end users due to their extensive, vague, and jargon-laden nature, replete with legal terminology that often requires a deeper level of specialized knowledge. The result is a lack of user-centric solutions to communicate privacy information transparently and understandably.

To bridge this gap, this thesis explores the concept of explainability as a crucial quality aspect for improving communication between systems and users concerning data practices, in a clear, understandable, and comprehensible manner. To this end, this thesis proposes an approach consisting of three theories supported by seven artifacts that outline the role of explainability in the context of privacy and provide guidelines for communicating privacy information. These theories and artifacts are intended to help software professionals (a) to identify privacy-relevant aspects, (b) to communicate them to users in a contextually relevant and understandable way, and (c) to design privacy-aware systems.

To validate the efficacy of the proposed approach, evaluations were conducted, including literature reviews, workshops, and user studies. The results endorse the suitability of the developed theories and artifacts, offering a promising foundation for developing privacy-aware, fair, and transparent systems.

Zusammenfassung

Im Zeitalter der fortschreitenden Digitalisierung, in dem die Technologie zunehmend in unsere Gesellschaft eindringt, rücken sogenannte *human values* wie Ethik, Fairness, Privatsphäre und Vertrauen weiter in den Mittelpunkt. Digitale Informationssysteme dringen immer stärker in private und berufliche Bereiche vor und bieten den Nutzern Unterstützung, schnell und einfach mit anderen Menschen in Kontakt zu treten, bei der Informationsbeschaffung und helfen bei der Erledigung täglicher Aufgaben. Im Gegenzug geben die Nutzer bereitwillig große Mengen an persönlichen Daten an diese Systeme weiter. Diese Datenerfassung bedeutet jedoch, dass die Privatsphäre der Nutzer zunehmend gefährdet ist. Daher ist die Aufklärung der Nutzer über die gesammelten Informationen und ihre anschließende Verarbeitung der Schlüssel, die Privatsphäre der Nutzer zu schützen.

Der Gesetzgeber hat Datenschutzerklärungen als Mittel zur Kommunikation von Datenpraktiken eingeführt. Leider erweisen sich diese Dokumente für die Endnutzer als praktisch nutzlos, da sie umfangreich, vage formuliert und mit Fachausdrücken gespickt sind, die oft ein tieferes Fachwissen erfordern. Das Ergebnis ist ein Mangel an nutzerorientierten Lösungen zur transparenten und verständlichen Vermittlung von Datenpraktiken.

Um diese Lücke zu schließen, wird in dieser Arbeit das Konzept der Erklärbarkeit als entscheidender Qualitätsaspekt zur Verbesserung der Kommunikation zwischen Systemen und Nutzern in Bezug auf Datenpraktiken in einer klaren, verständlichen und nachvollziehbaren Weise untersucht. Zu diesem Zweck wird ein Ansatz vorgeschlagen, der aus drei Theorien besteht, die durch sieben Artefakte gestützt werden, die die Rolle der Erklärbarkeit im Kontext der Privatsphäre skizzieren und Leitlinien für die Kommunikation von Datenschutzinformationen aufstellen. Diese Theorien und Artefakte sollen Software-Experten unterstützen, (a) privatsphärerelevante Aspekte zu identifizieren, (b) diese kontextrelevant und verständlich an den Nutzer zu kommunizieren, um (c) datenschutzfreundliche Systeme zu designen. Um die Wirksamkeit des vorgeschlagenen Ansatzes zu validieren, wurden Evaluierungen durchgeführt, darunter Literaturrecherchen, Workshops und Nutzerstudien. Die Ergebnisse bestätigen die Eignung der entwickelten Theorien und Artefakte und bieten eine vielversprechende Grundlage für die Entwicklung datenschutzfreundlicher, fairer und transparenter Systeme.

1 Einführung

Wir leben im digitalen Zeitalter, in dem wir allgegenwärtig von intelligenten, digitalen Geräten wie Smartphones, Smartwatches, etc. umgeben sind [11]. Diese Geräte sind so eng mit unserem Alltag verwoben, das wir sie kaum noch wahrnehmen. Die auf diesen Geräten laufenden Algorithmen unterstützen uns beispielsweise bei der Entscheidungsfindung [12, 13, 14], formen und beeinflussen somit wiederum unsere Gesellschaft [12, 15, 16]. So helfen uns Algorithmen im Gesundheitswesen bei der Diagnose von Krankheiten [17, 18], bei der Vermittlung von Mitfahrgelegenheiten [19] oder auch in der Justiz, um beispielsweise Rückfälligkeit bei straffällig gewordenen Personen vorherzusagen [20]. Es gibt aber auch invasivere Arten, wie Algorithmen Einflüsse auf die Gesellschaft nehmen. Sie können Meinungen in der Gesellschaft beeinflussen oder verzerren [21] und schließlich sogar Einfluss auf Wahlen nehmen [22, 23, 24]. Der *Treibstoff* dieser Algorithmen sind dabei Daten.

1.1 Motivation

Im Jahr 2017 titelte der *The Economist*, „die wertvollste Ressource der Welt ist nicht mehr Öl, sondern Daten“ [25, 26]. Persönliche Daten sind längst zu einer Art virtuellen Währung geworden [27, 28, 29], wodurch eine florierende Industrie entstehen konnte. Die sogenannten *Data*

Broker. Bei jeder Interaktion mit einem Informationssystem hinterlassen wir unseren digitalen Fußabdruck, denn all die dabei entstehenden Daten werden für gewöhnlich gespeichert, zusammengeführt und ausgewertet [26, 30, 31]. Wenn wir im Auto fahren, wird beispielsweise der Druck, mit dem wir auf das Gaspedal treten sowie auch die damit verbundene Geschwindigkeit gespeichert. Diese Daten können anschließend Diagnosezwecken dienen oder bei der Untersuchung von Unfällen herangezogen werden [31].

Mit jeder Interaktion von Software-Systemen gehen wir also bewusst oder unbewusst einen *Trade-Off* in Bezug auf die Vorteile ein, die die Nutzung des Systems uns bietet sowie den persönlichen Daten, die wir dabei hinterlassen [32]. Gründe für diesen impliziten Datenaustausch sind auf Nutzerseite motiviert durch den Konsum von „freien" Informationen [33], personalisierten Inhalten, Rabatte [34], Treueprogramme [35] und andere ökonomische Anreize [36]. Obwohl diese persönlichen Daten uns gehören, geschieht die Offenlegung dieser für gewöhnlich ohne unsere explizite Zustimmung und oft auch ohne unser Wissen [37]. Im Englischen spricht man von „*informed consent*", was übersetzt „informierte Zustimmung" oder „Einwilligung nach Aufklärung" bedeutet. Beide Übersetzungen sind im deutschen jedoch nicht eindeutig und in diesem Kontext eher ungebräuchlich. Daher verwende ich im Rahmen meiner Dissertation den Terminus der „expliziten Zustimmung". Dieser beinhaltet im Kontext der Privatsphäre die Aufklärung sowie ein Verständnis über die Nutzung der eigenen persönlichen Daten **und** die damit einhergehende Zustimmung dieser Nutzung.

Viele Internetseiten (über 60% in Europa [38]) setzen auf sogenannte *Cookie*[1] *Banner* (Abschnitt 2.6.3.1), im Englischen zutreffender als *Cookie Consent Notice* bezeichnet. Sie sollen dazu dienen, Seitenbesucher über die Datenpraktiken eines Anbieters zu informieren und anschließend dessen Zustimmung für die Datennutzung einzuholen. Allerdings weisen die Implementationen dieser Cookie Banner meist substantielle Usability-Probleme auf [38, 40]. Demzufolge ist es häufig für den Seitenbesucher unklar, welche Daten gesammelt sowie gespeichert werden und vor allem – zu welchem Zweck. Neben den Cookie Bannern gibt es zusätzlich noch die Datenschutzerklärungen (DSEs). DSEs oder auch kurze Datenschutzhinweise (englisch: *privacy notices*) stellen den primären Kanal dar, über den Service-Anbieter wie Internetseitenbetreiber, Social Media-Anbieter oder Streaming-Dienste, aber auch Banken, Versicherer, Telekommunikationsanbieter etc. ihre Konsumenten über die eigenen Datenpraktiken informieren. Das Problem bei dieser Art Medium ist, dass Benutzer sie weitestgehend ignorieren [41, 42, 43, 44], da DSEs aus sehr umfangreichen, schwer zu verstehenden Texten bestehen, die vage und verwirrend formuliert sind sowie juristisches Hintergrundwissen für

[1] „Zeichenfolge, die mit einer Web-Seite vom Server geladen werden kann und bei einer erneuten Anfrage an den Server mitgesendet wird. Sinn ist, unter anderem Besucher wiederzuerkennen, so dass es beispielsweise nicht erforderlich ist, Nutzerdaten neu einzugeben" [39]

tiefer gehendes Verständnis erfordern [45, 46, 47, 48, 49], denn sie sind „von Anwälten für Anwälte“ geschrieben [41, 49]. Somit stellen sowohl DSEs als auch Cookie Banner kein geeignetes Medium zur Aufklärung der Nutzer über Datenpraktiken dar.

Dieses Dilemma macht deutlich, wie wichtig es ist, über Alternativen bei der Aufklärung über die Verwendung unserer persönlichen Daten nachzudenken und diese zu erforschen. Der Ansatz sollte hierbei nutzerzentriert sein und sich zum Ziel setzen, mehr Transparenz, Vertrauenswürdigkeit und ethischen Umgang mit den Daten zu fördern. Bezugnehmend auf Garcia-Rivadulla [32] sollten Unternehmen dies nicht als Bedrohung wahrnehmen, sondern als „eine Chance, im Zusammenhang mit dem Schutz der Privatsphäre innovativ zu sein und das Vertrauen der Verbraucher zu gewinnen“ [41]. Cummings et al. [50] zeigten in einer Studie, wenn Nutzern verständliche Informationen über Datenpraktiken bereitgestellt werden, sind diese eher gewillt, ihre persönlichen Daten preiszugeben. Allerdings stellten Phelps et al. [51] auch fest, dass es je nach Art der Information deutlich variieren kann. Zusammenfassend führt dieses Dilemma nun zu der folgenden Problemstellung für diese Dissertation:

Problemstellung Während Datenschutzerklärungen (DSEs) oder auch kurze Datenschutzhinweise als primärer Kanal verwendet werden, um Hinweise über Datenpraktiken eines Anbieters oder Services zu transportieren und zu vermitteln, bietet diese Art der *statischen* Dokumente wenig aufklärende und verständliche Informationen für Benutzer. Vielmehr sollte eine nutzerzentrierte und -orientierte Lösung angestrebt werden. Hierfür könnte das Konzept der Erklärbarkeit als eine Art Hebel betrachtet werden bzw. Mittel zum Zweck sein. Denn Erklärbarkeit als nicht-funktionale Anforderung (NFR) ermöglicht mangelnde Transparenz eines Systems abzuschwächen und den Endnutzern ein Verständnis für das Verhalten eines Systems zu vermitteln, indem Erklärungen gegeben und Informationen offengelegt werden [4, 52]. Das wiederum bedeutet, dass andere benutzerorientierte Qualitätsziele durch die Spezifikation und Erfüllung von Erklärbarkeitsanforderungen erfüllt werden können. Es ist daher notwendig zu erforschen, wie die Erklärbarkeit hier die Informationslücke zwischen Anwender und zur Anwendung kommenden Datenpraktiken schließen kann, in dem die für das Verstehen und Nachvollziehen notwendigen Informationen angemessen und zufriedenstellend transportiert werden. Darüber hinaus sollten Konzepte entwickelt werden, wie solche Lösungen in Software-Systeme eingebunden werden können und mit Hilfe des Potentials der Erklärbarkeit die Privatsphäre der Nutzer besser respektiert werden kann.

1.2 Forschungsziel

Basierend auf der Problemstellung verfolgt die Doktorarbeit das nachfolgend formulierte Forschungsziel, um Möglichkeiten zu untersuchen, Endbenutzer mit verständlichen, nachvollziehbaren und zweckdienlichen Informationen bezüglich ihrer Privatsphäre zu versorgen. Dazu habe ich mich bei der Formulierung des Forschungsziels an der Goal Definition-Vorlage [53, 54] orientiert, um sicherzustellen, dass der Forschungsrahmen dieser Doktorarbeit klar definiert und eindeutig abgesteckt ist.

Forschungsziel **Analysiere** die Nutzung des Konzepts der Erklärbarkeit zur Kommunikation von Privatsphäreaspekten **zum Zweck der** Informierung und Aufklärung der Systemnutzer **in Bezug auf** Usability und Effektivität der kommunizierten Informationen **aus der Sicht von** Endbenutzern **im Kontext von** Software-Systemen, die private Daten von Nutzern sammeln und verwenden.

Das Ziel dieser Dissertation ist zu untersuchen, ob sich das Konzept der Erklärbarkeit eignet, um mit Endbenutzern in einen Dialog über die Nutzung ihrer persönlichen Daten zu treten. Hierbei stehen die Aspekte *Usability* und *Effektivität* im Fokus. Beide beziehen sich auf die kommunizierten Informationen aus Sicht der Benutzer, also deren Wahrnehmungen bezüglich der Erklärungen. Nielsen [55] sagt, „Usability bezieht sich auf alle Aspekte eines Systems, mit denen ein Mensch interagieren kann, einschließlich der Installations- und Wartungsprozeduren“. Demzufolge würde auch ein *effektiver* Umgang mit einem System in den Bereich der Usability fallen. Ich trenne aber bewusst bei der Formulierung des Forschungsziels diese zwei Begrifflichkeiten auf, um beide Aspekte abgrenzend voneinander betrachten zu können. Somit bezieht sich der Aspekt der Usability in diesem Kontext auf den technischen Rahmen eines Software-Systems. Also wie ist die Erklärung implementiert, stört diese den Benutzer oder behindert sie ihn sogar bei der Nutzung des Systems. Usability beinhaltet im Software Engineering (SE) mehr als nur rein technische Aspekte.

Die Effektivität konzentriert sich auf Verständnis, Nachvollziehbarkeit und Akzeptanz der kommunizierten Informationen. Hat der Benutzer verstanden, dass (a) Daten genutzt werden, (b) welche Daten das sind sowie (c) zu welchem Zweck, ist er somit in der Lage, der Nutzung explizit zustimmen zu können. Zusammenfassend bezeichnet die Usability also die technische Umsetzung von Erklärungen und die Effektivität die Güte der transportierten Informationen, beides aus Sicht (Wahrnehmung) der Endbenutzer. Basierend auf dem Forschungsziel ergeben sich nun folgende Forschungsfragen (RQs):

***RQ1:* In welcher Beziehung stehen Erklärbarkeit und Privatsphäre zueinander und wie ist deren Auswirkung auf die Softwarequalität?**

Erklärbarkeit (englisch: *explainability*) und Privatsphäre (englisch: *privacy*) sind hier als Qualitätsaspekte[2] von Software-Systemen zu verstehen. Es geht bei der Untersuchung dieser Frage darum, welche Wechselbeziehung es zwischen diesen beiden Qualitäten gibt und welcher Art diese Beziehung ist. Dieses Verständnis ist notwendig, um effektive Lösungen zu erforschen, wie das Konzept der Erklärbarkeit im Kontext der Privatsphäre sinnvoll und zielführend eingespannt werden kann. Dafür ist es darüber hinaus wichtig, die Auswirkungen beider auf die Softwarequalität zu verstehen. Dazu ist ein detaillierter Überblick notwendig, welche Qualitätsaspekte durch Erklärbarkeit und Privatsphäre beeinflusst werden können.

***RQ2:* Wie kann die Verwendung von Privatsphäreaspekten erklärt werden?**

Dokumente wie z.B. DSEs sind offensichtlich keine geeignete Art, Nutzer über die Verwendung ihrer persönlichen Daten zu informieren, respektive Sie aufzuklären, um diese wiederum in die Lage zu versetzen, bewusste Entscheidungen über die explizite Zustimmung ihrer Daten treffen zu können. Daher soll, auch auf Grundlage der Ergebnisse aus RQ1, untersucht werden, wie das Konzept der Erklärbarkeit dazu beitragen kann, Endbenutzer hinsichtlich der Nutzung ihrer persönlichen Daten zu informieren und aufzuklären und zwar mit dem Ziel, dass diese schlussendlich bewusste Entscheidungen in Bezug auf ihre Privatsphäre treffen können.

***RQ3:* Welchen Einfluss haben Erklärungen von Privatsphäreaspekten auf den Endbenutzer und was ist bei deren Umsetzung zu beachten?**

Damit ein System verwendet wird, muss es von seinen Nutzern akzeptiert werden. Dies beinhaltet nach Nielsen [55] die *practical acceptance* wie auch die *social acceptance*. Laut Nielsen umfasst die practical acceptance Aspekte wie Kosten, Kompatibilität und Zuverlässigkeit, aber auch die Nützlichkeit (englisch: *usefulness*). Bei Letzterer geht es laut Nielsen um die Frage, „ob das System genutzt werden kann, um ein gewünschtes Ziel zu erreichen“ [55]. Die Social acceptance beinhaltet u.a. Aspekte wie Transparenz, Fairness und Ethik [56]. Die Frage zielt nun darauf ab, die praktische Akzeptanz von Erklärungen zu Privatsphäreaspekten zu untersuchen wie auch deren sozialer Akzeptanz, was auch Verständnis

[2] Der Begriff der Qualitätsaspekte wird in Abschnitt 2.3 erläutert.

und Nachvollziehbarkeit umfasst, ebenso Design-Konzepte zu untersuchen, die Akzeptanz und Verständnis auf Seiten der Endbenutzer ermöglichen und fördern.

1.3 Forschungsansatz

Um die Forschungsfragen zu beantworten und somit das Forschungsziel zu erreichen, orientiere ich mich am Vorgehen von Design Science und Information System Research. Beim Design Science geht es um die Schaffung von Theorien und Artefakten, die ein zuvor identifiziertes relevantes organisatorisches Informationstechnologie (IT) Problem adressieren und lösen [57]. Zudem wird dieses geschaffene Artefakt (oder auch die Theorie) anhand einer vorhanden bzw. geschaffenen Wissensbasis evaluiert. Weiterführende Details zu Design Sciene und Information System Research sind im Kapitel 3 zu finden. Folgende Artefakte liefert die vorliegende Dissertation: Definitionen für erklärbare Systeme, Privatsphäre und Online-Privatsphäre, Modell der Beziehung von Erklärbarkeit und Privatsphäre im Zusammenspiel mit anderen Qualitätsaspekten, das Konzept der Erklärungen zu Privatsphäreaspekten, Anforderungen für deren Umsetzung, Einflussfaktoren, Design-Empfehlungen und den Vorschlag des Trialogs der Privatsphäre.

1.4 Struktur der Arbeit

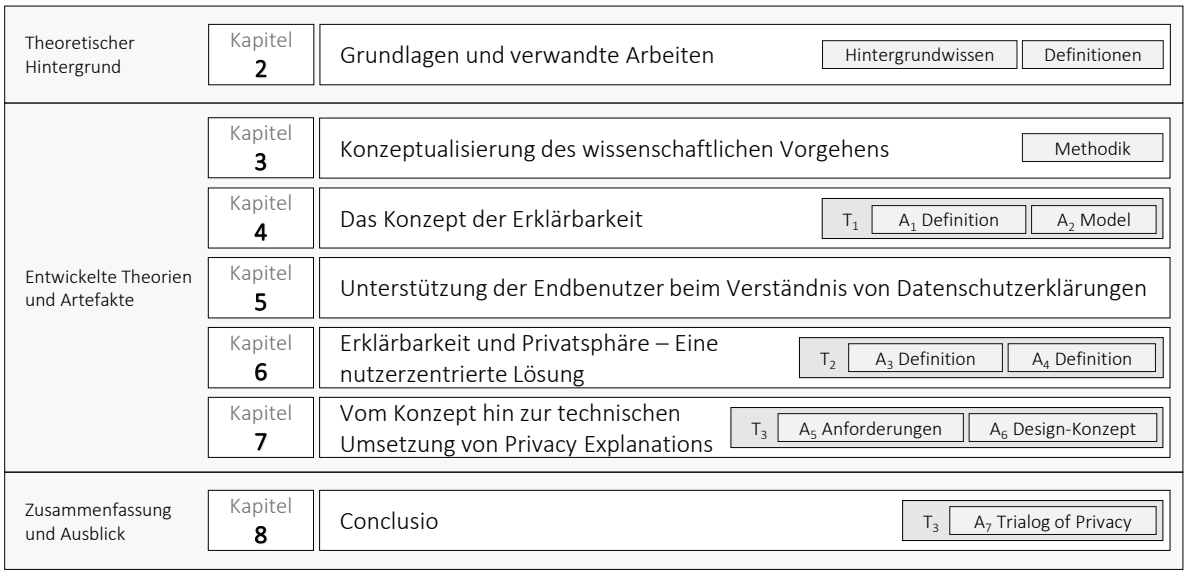


Abbildung 1.1: Struktur der Dissertation. Legende: T_i: Theorie, A_i: Artefakt

Die Struktur der Doktorarbeit ist in Abbildung 1.1 abgebildet. Zunächst vermittelt Kapitel 2 das zum weiteren Verständnis nötige Hintergrundwissen und stellt verwandte Arbeiten

vor. Kapitel 3 beleuchtet das Forschungsvorgehen im Detail. Das Konzept der Erklärbarkeit wird in Kapitel 4 diskutiert und bildet die Forschungsbasis für diese Dissertation. Kapitel 4 liefert zudem die ersten beiden Forschungsartefakte A_1 und A_2[3]. Ein erster Ansatz hin zu Erklärungen für die Privatsphäre wird in Kapitel 5 untersucht. Chronologisch und aus Sicht meiner Forschung darauf aufbauend beschäftigt sich Kapitel 6 mit einer Verfeinerung der Erkenntnisse aus den vorangegangenen Kapiteln und mündet in dem Konzept der Privacy Explanations[4](A_3), dessen prototypische Umsetzung Schritt für Schritt in Kapitel 7 diskutiert wird. Kapitel 8 schließt mit einer Zusammenfassung und einem Ausblick.

[3] Details zu den Artefakten sind in Abschnitt 3.1 zu finden

[4] Für eine Definition des Begriffs „Privacy Explanation“ siehe Abschnitt 6.1.3

2

Grundlagen und verwandte Arbeiten

2.1 Grundlagen

Diese Dissertation beschäftigt sich mit dem Thema Datenschutz und Privatsphäre im Kontext des Software bzw. Requirements Engineerings. In diesem Kapitel vermittle ich Grundlagen, gebe wichtige Hintergrundinformationen und definiere Begriffe, die für das weitere Verständnis meiner Dissertation notwendig sind.

Laut Valacich [58] ist ein Informationssystem[1] „die Kombination aus Menschen und Informationstechnologie, die nützliche Daten erstellen, sammeln, verarbeiten, speichern und verteilen“. Für meine Dissertation verschmelze ich Valacichs [58] Auslegung des Begriffs *Informationssystem* und Sommervilles [59] Begriffsbestimmung für *Software-Systeme* zur nachfolgenden Definition, denn Software-Systeme lassen sich als Untermenge von Informationssystemen verstehen:

[1] Im Rahmen meiner Dissertation verwende ich die Begriffe System und Informationssystem synonym.

Definition 2.1.1: Informationssystem [58, 59]

Informationssysteme bestehen aus Hardware, Software (-Systemen) und Telekommunikationsnetzen, die von Menschen entwickelt, gemanagt und genutzt werden, um zweckdienliche Daten zu sammeln, zu erstellen und zu verbreiten, häufig in einem Unternehmenskontext. Ein **Software-System** besteht aus einer Reihe von separaten Computer-Programmen sowie zugehörigen Konfigurationsdateien und Benutzerdokumentationen, die zusammen operieren und wiederum ein Informationssystem formen.

Ein weiterer wichtiger Terminus ist der des *Stakeholders*. Bezugnehmend auf den IEEE[2]-Standard 1471 [60] ist ein Stakeholder „ein Individuum, ein Team oder eine Organisation (oder Klassen davon) mit Interessen oder Anliegen [(englisch: *stakes*)] in Bezug auf ein [Software-]System". Laut Freeman [61] umfasst der Begriff des Stakeholders jede Person, Gruppe oder Organisation, die die Erreichung der Ziele der Organisation oder des untersuchten Systems beeinflussen kann oder davon betroffen ist. Dieser Einfluss kann sowohl negativ als auch positiv sein und wirkt sich auch auf die Anforderungen[3] aus [61, 62, 63, 64]. Darüber hinaus clustert McManus [65] Stakeholder als primäre (Eigentümer, Lieferanten, Mitarbeiter, Kunden, lokale Gemeinschaften und Manager), sekundäre, erweiterte und externe Stakeholder. Es ist wichtig, nicht nur die von einem System positiv betroffenen Akteure zu berücksichtigen. Stattdessen ist es auch wichtig, die *negativen Stakeholder* zu berücksichtigen und zu identifizieren [66, 67]. Ein gutes Beispiel für einen negativen Stakeholder geben Ballejos et al. [66]. Sie definieren sie als „diejenigen, die als Folge der Systemimplementierung eine Art von Schaden erleiden oder durch die Entwicklung des Systems beeinträchtigt werden (z.B. Verlust des Arbeitsplatzes, Verlust der Entscheidungsbefugnis, physischer Schaden, finanzieller Schaden usw.)" [66, p.285]. Auf der Grundlage der Gemeinsamkeiten dieser Definitionen definiere ich Stakeholder wie folgt:

Definition 2.1.2: Stakeholder

Ein **Stakeholder** ist eine Person oder eine Gruppe von Personen, die ein Anliegen (englisch: *stake*) an einem System haben (in positiver oder negativer Weise) und/oder direkt oder indirekt von einem System betroffen sind.

Darüber hinaus übernehme ich die Definition für Software Engineering (SE) von Sommerville [59]:

[2] IEEE steht für Institute of Electrical and Electronics Engineers
[3] Der Begriff der Anforderung wird in Definition 2.2.1 definiert.

Definition 2.1.3: Software Engineering [59]

Das **Software Engineering** befasst sich mit allen Aspekten der Software-Produktion, von den frühen Phasen der Systemspezifikation bis zur Wartung des Systems nach seiner Inbetriebnahme.

Ebenfalls basierend auf Sommerville [59] definiere ich den Software Engineer[4] wie folgt:

Definition 2.1.4: Software Engineer [59]

Ein **Software Engineer** ist eine Person, die sich mit der Spezifikation, dem Design, der Konstruktion, der Entwicklung, dem Vertrieb (Rollout) und der Wartung von Software-Systemen beschäftigt. Beispiele für Rollen des Software Engineers sind: Requirements Engineers, Software-Architekten, Entwickler und Tester.

2.2 Requirements Engineering

Requirements Engineering (RE) ist ein Feld des SE und beschäftigt „sich mit den realen Zielen für die Funktionen von Software-Systemen und den Beschränkungen“ [68]. Der Erfolg eines solchen Systems wird daran gemessen, „inwieweit es den Zweck [(reale Ziele)] erfüllt, für den es gedacht war“ [69]. Nuseibeh und Easterbrook geben an, dass RE der Prozess ist, der die „Entdeckung dieses Zwecks durch die Identifizierung der Beteiligten [Stakeholder] und ihrer Bedürfnisse und die Dokumentation dieser Bedürfnisse in einer Form, die eine Analyse, Kommunikation und anschließende Implementierung ermöglicht“ [69]. Dokumentiert werden diese Bedürfnisse mit Hilfe sogenannter Anforderungen (englisch: *requirements*). Für meine Dissertation nutze ich die Definition von Glinz [70], die die offizielle Definition des *International Requirements Engineering Board* (IREB) darstellt:

Definition 2.2.1: Anforderung (Requirement) [70]

1. Ein von einem Stakeholder wahrgenommener Bedarf.
2. Eine Fähigkeit oder Eigenschaft, die ein System haben muss.
3. Eine dokumentierte Darstellung eines Bedarfs, einer Fähigkeit oder einer Eigenschaft.

Für das RE existieren eine Reihe unterschiedlicher Definitionen, Referenzmodelle oder auch Frameworks, die einen Überblick über das RE geben [62, 68, 69, 71, 72, 73, 70, 74, 1].

[4] Ich behalte den englischen Begriff des Engineers bei, da die deutsche Rolle des Ingenieurs hier nicht ganz zutreffend ist.

Alle beschreiben im Kern aber die gleichen Ziele und weisen eine gemeinsame Basis auf. Im Rahmen meiner Dissertation verwende ich die folgende Definition von Glinz [70]:

Definition 2.2.2: Requirements Engineering [70]

Ein systematischer und disziplinierter Ansatz für die Spezifikation und das Management von Anforderungen mit den folgenden Zielen:
(1.) Kenntnis der relevanten Anforderungen, Erzielung eines Konsenses zwischen den Stakeholdern über diese Anforderungen, Dokumentation nach vorgegebenen Standards und systematische Verwaltung dieser Anforderungen,
(2.) Verstehen und Dokumentieren der Wünsche und Bedürfnisse der Stakeholder,
(3.) Spezifizieren und Verwalten von Anforderungen, um das Risiko zu minimieren, ein System zu liefern, das nicht den Wünschen und Bedürfnissen der Stakeholder entspricht.

In Abbildung 2.1 ist das von Börger et al. [1] vorgeschlagene RE-Referenzmodell dargestellt, welches das Requirements Engineering in zwei Hauptbereiche mit den jeweils damit verknüpften Aktivitäten untergliedert. Das **Requirements Management** umfasst das *Change Management* und *Tracing*. Diese Aktivitäten befassen sich mit der Verwaltung von ausgearbeiteten Anforderungen und damit verbundenen Informationen, um *Change Requests* und deren Nachverfolgbarkeit zu gewährleisten. Im Rahmen dieser Dissertation liegt der Schwerpunkt eher bei der Requirements Analysis und den damit verbundenen Aktivitäten, also mit Fokus auf den Bedürfnissen und Erwartungen von Endbenutzern an ein Informationssystem. Das Requirements Management sei daher nur der Vollständigkeit halber erwähnt.

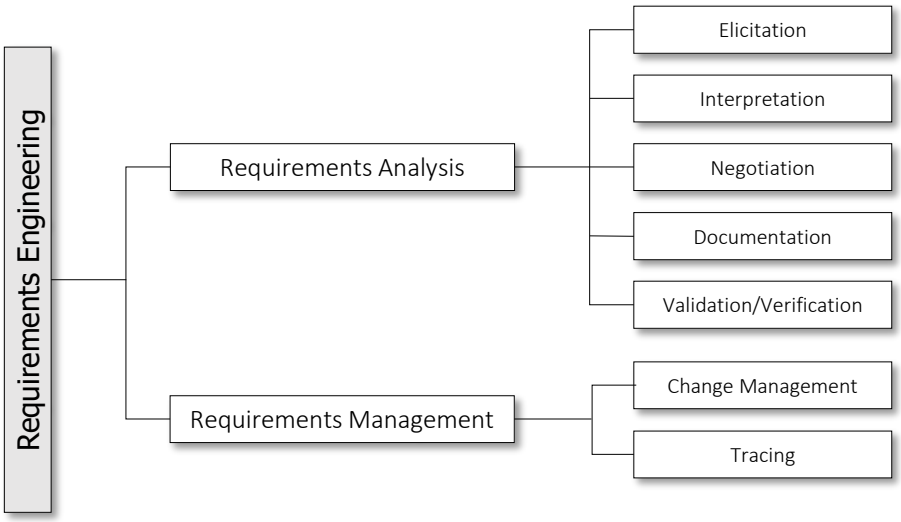


Abbildung 2.1: Requirements Engineering Referenzmodell nach [1]

Die **Requirements Analysis** beinhaltet die Aktivitäten, die sich mit der Exploration, Bewertung, Dokumentation und schließlich der Bestätigung von Anforderungen an ein Informationssystem befassen. Ein Großteil der Fehler und Risiken, die bei der Entwicklung von Systemen auftreten, resultieren aus unzureichend definierten Anforderungen [73, 75, 76] und entstehen in der Phase der Requirements Analysis, was wiederum hohe Kosten verursacht [73]. Daher ist diese Phase von entscheidender Bedeutung für den Erfolg eines Systems.

Eliciation Aufgabe der Elizitierung ist, relevante Informationen zu gewinnen sowie Quellen für Anforderungen zu identifizieren wie z.B. Bedürfnisse und Ziele der Stakeholder, deren Aufgaben sowie deren (Arbeits-) Umgebungen. Aber auch Altsysteme oder existierende Dokumente können solche Quelle darstellen. Ein Requirements Engineer muss hierfür eng mit den Stakeholdern zusammenarbeiten, damit er auf der Grundlage der gewonnen Informationen die Grenzen des zu entwickelnden Systems definieren kann. Mit Hilfe dieses definierten *Scope* des Systems, kann der Requirements Engineer nun Bedürfnisse und (Qualitäts-) Erwartungen der Stakeholder in sogenannten *Rohanforderungen* (abstrakte Anforderungen) festhalten [73].

Interpretation Die Interpretation dient dazu, ein tieferes Verständnis über die erhobenen Rohanforderungen zu erhalten. Diese werden während der Interpretation analysiert und strukturiert, wie z.B. in funktionale und nichtfunktionale Anforderungen, Geschäftsziele, Aufgaben usw. So können Abhängigkeiten, Konflikte, Unklarheiten und mögliche Lücken aufgedeckt werden, die im nächsten Schritt (Negotiation) dann geklärt werden können.

Negotiation Im Allgemeinen muss ein System die unterschiedlichen Bedürfnisse, Wünsche und Ziele verschiedener Stakeholder erfüllen. Jeder Stakeholder bringt hier seine ganz eigene, persönliche Sicht und Erfahrung mit, so dass es zu unterschiedlichen Meinungen und Konflikten kommen kann. Diese sollen in dieser Aktivität sichtbar gemacht werden. Requirements Engineer und Stakeholder arbeiten hierfür dann Kompromisse und Lösungen aus.

Documentation Der Schwerpunkt dieser Aktivität „liegt in der Dokumentation und Spezifikation der erhobenen Anforderungen nach den festgelegten Dokumentations- und Spezifikationsregeln“ [71]. Dieser Schritt dient dazu, das zuvor gewonnene Wissen rund um die Anforderungen dauerhaft zu speichern und organisiert darzustellen.

Validation/Verification Validierung und Verifikation dienen der formalen und inhaltlichen Prüfung der elizitierten Anforderungen. Die Validierung beschreibt den Prozess der *Bestätigung* der Anforderungen. Es wird geprüft, ob die Anforderungen mit den Bedürfnissen der Stakeholder übereinstimmen. Bei der Verifikation wird geprüft, ob die Anforderungen auch korrekt spezifiziert wurden. Es wird also geschaut, ob erhobene Rohanforderungen und zugehörige Informationen auch mit den spezifizierten Anforderungen übereinstimmen.

2.2.1 Gemeinsames Verständnis – Shared Understanding

RE umfasst nicht nur das Erheben und Spezifizieren von Anforderungen, sondern erfordert auch eine effektive Kommunikation mit und zwischen den verschiedenen Stakeholdern [69, 77]. Ein gemeinsames Verständnis (englisch: *shared understanding*) spielt hier eine ganz entscheidende Rolle für den Erfolg von qualitativ hochwertiger Software, die den Bedürfnissen der Stakeholder gerecht werden soll [78, 79]. Vereinfacht gesagt, kann gemeinsames Verständnis erreicht werden, wenn alle Beteiligten ein äquivalentes mentales Modell haben. Bezugnehmend auf Norman sind mentale Modelle „konzeptionelle Modelle in den Köpfen der Menschen, die ihr Verständnis davon darstellen, wie Dinge funktionieren. Verschiedene Menschen können unterschiedliche mentale Modelle desselben Gegenstandes haben. Ein und dieselbe Person kann sogar mehrere Modelle desselben Gegenstands haben, von denen sich jedes mit einem anderen Aspekt seiner Funktionsweise befasst: Die Modelle können sogar in Konflikt miteinander stehen“ [80, S. 26]. Diese „Dinge“ oder „Gegenstände“ können auch Informationssysteme oder Teile davon sein. Ich definiere den Begriff des mentalen Modells in Anlehnung an Norman [80] wie folgt:

Definition 2.2.3: Mentales Modell

Ein mentales Modell ist die konzeptionelle Vorstellung eines Individuums, die dessen individuelles Verständnis davon darstellt, wie etwas (beispielsweise ein System oder Teile davon) funktioniert.

Nachdem nun der Begriff des mentalen Modells definiert ist, definiere ich den Begriff des gemeinsamen Verständnisses unter Berufung auf Easterbrook [81] wie folgt:

Definition 2.2.4: Gemeinsames Verständnis (Shared Understanding)

Zwei oder mehr Personen teilen ein gemeinsames Verständnis einer Situation, wenn die Elemente ihrer mentalen Modelle, die sich auf diese Situation beziehen, im Einklang stehen. Im Einklang stehen bedeutet, dass ihre mentalen Modelle dieselben Erklärungen, Erwartungen und Vorhersagen für diese Situation liefern.

2.3 Softwarequalität

„An Softwarequalität denkt man im normalen Leben eigentlich nur, wenn etwas damit nicht stimmt“ [82]. Probleme bei der Qualität von Software können unterschiedliche Auswirkungen haben, wie zum Beispiel das Umgehen eines Kopierschutzes bei einem Videospiel [83],

das Offenlegen privater Daten[5] oder auch katastrophale Auswirkungen wie das böswillige Eindringen in das Kontrollsystem eines Kernkraftwerks [83]. Softwarequalität ist ein mehrdimensionales Konzept [85], steht im direktem Bezug zur Kundenzufriedenheit [86] und Gebrauchstauglichkeit[6] [88]. Der *IEEE Standard for Software Quality Assurance Processes (IEEE Std. 730-2014)* sagt, Softwarequalität ist „der Grad, in dem ein Softwareprodukt festgelegte Anforderungen erfüllt; die Qualität hängt jedoch davon ab, inwieweit diese festgelegten Anforderungen die Bedürfnisse, Wünsche und Erwartungen der Stakeholder genau wiedergeben“ [89]. Beschrieben werden derartige Merkmale der Qualität u.a. durch nicht-funktionale Anforderungen (NFRs) [90]. Im Kontext meiner Dissertation verwende ich den Begriff des Qualitätsaspekts und synonym auch Qualitätsmerkmals, um mich sowohl auf NFRs als auch auf Aspekte zu beziehen, die sich wiederum auf NFRs beziehen oder diese zusammensetzen. In diesem Sinne folge ich Glinz und betrachte NFRs als „ein Attribut oder eine Einschränkung eines Systems“ [91].

2.4 Erklärbarkeit

Häufig ist nicht ganz klar, warum ein Software-System ein bestimmtes Ergebnis anzeigt oder warum eine bestimmte Interaktion mit einem System nicht zum gewünschten Ziel führt. Das kann wiederum zu Frustration auf Seiten des Benutzers führen [92]. Erklärbarkeit kann dazu beitragen, derartige Irritationen abzuschwächen und Benutzern helfen, Verhalten oder Ergebnisse von Systemen besser zu verstehen. Erklärbarkeit ist kein neues Konzept und ist häufig in der Domäne Maschinelles Lernen (ML) bzw. genauer gesagt bei eXplainable Artificial Intelligence (XAI) anzutreffen und wird dort ebenfalls intensiv beforscht. In der Literatur wird Erklärbarkeit häufig mit den Begriffen Interpretierbarkeit, Transparenz und Verständlichkeit (englisch: *understandability*) gleichgesetzt. In der Tat sind diese Begriffe und die zugrunde liegenden Konzepte dahinter eng miteinander verwoben, ich messe der Erklärbarkeit aber eine eigenständige Rolle bei. „Genauer gesagt sind Erklärungen Operationalisierungen der Erklärbarkeit (objektiver Faktor). Erklärungen können durch die Bereitstellung von Informationen die Interpretierbarkeit und Verständlichkeit des Systems beeinflussen (subjektive Faktoren)“ [56]. Im Allgemeinen haben Erklärungen ein Zweck zu erfüllen. Aus philosophischer Sicht ist eine Erklärung eine Antwort auf eine vorgegebene Frage, warum etwas so ist oder passiert [93, 94]. Darüber hinaus können Erklärungen auch als Argumente definiert werden, die darlegen, wie das zu Erklärende (*Explanandum*) eine Art der Ableitung aus Naturgesetzen und empirischen Bedingungen ist [95]. Diesem Gedankengang folgend, füllt eine

[5] Durch eine Sicherheitslücke im sozialen Netzwerk Google+ wurden die privaten Daten von fast 500000 Personen offengelegt, die das soziale Netzwerk zwischen 2015 und März 2018 genutzt haben [84].

[6] Im englischen Original: „fitness for purpose“ [87].

Erklärung also eine mögliche Wissenslücke und kann zum Verständnis beitragen, indem sie Zustände der Verwirrung oder auch Abweichung adressiert, die auftreten können, wenn das verinnerlichte mentale Modell eines Individuums im Konflikt mit einer bestimmten Situation steht, was auch von Schank [96] so beschrieben wird. Das bedeutet, dass Erklärungen es ermöglichen, das mentale Modell eines Benutzers anzugleichen und ggf. auch zu korrigieren. Für meine Dissertation nutze ich sowohl die Definition einer Erklärung (englisch: *explanation*) also auch die Definition von Erklärbarkeit (englisch: *explainability*) von Chazette [56].

Definition 2.4.1: Erklärung und Erklärbarkeit nach [56]

Eine **Erklärung** ist eine Information, die dazu beiträgt, dass der Adressat ein Explanandum in einem bestimmten Kontext versteht.

Erklärbarkeit ist die Fähigkeit oder der Akt der Offenlegung von Informationen, die für einen Adressaten notwendig sind, um einen bestimmten Aspekt eines Systems in einem bestimmten Kontext zu verstehen, was durch die Bereitstellung von Erklärungen erreicht werden kann.

2.5 Privatsphäre

Einige Autoren differenzieren zwischen den Begriffen *Privatsphäre* und *Privatheit* [97, 98, 99]. So beschreiben Behrendt et al. [99] das das Wort „Privatheit“ ein aus dem englischen Wort „Privacy“ abgeleiteter Neologismus ist. „Näher an der deutschen Alltagssprache [ist] der der Ausdruck ‚Privatsphäre‘“ [99, vgl.]. Allerdings – und damit beziehen sich die Autoren auf das in *Privatsphäre* enthaltene Wort *Sphäre* – sind sie der Ansicht, das Privat*sphäre* „stark auf einen geschützten Ort rekurriert, wohingegen das Private in seiner gesamten Bedeutungsbreite daneben auch dezisionale und informationelle Bereiche umfasst, für die die Ortsmetapher unpassend ist“ [99]. Semantisch gesehen stimme ich hiermit überein, entscheide mich aber dennoch im Rahmen meiner Dissertation für den Begriff der *Privatsphäre*, da dieser gebräuchlicher ist und somit das Gesamtverständnis vereinfacht.

Privatsphäre ist ein normatives Konzept. Es ist nichts Neues für unsere moderne Gesellschaft, sondern existiert schon seit sehr langer Zeit und ist tief in soziologischen, philosophischen, rechtlichen, politischen und wirtschaftlichen Traditionen verwurzelt [100]. In der Vergangenheit beispielsweise, als die Menschen noch Jäger und Sammler waren, war es von entscheidender Bedeutung zu wissen, wo und wann es reife Früchte gab oder wo sich die nächste Wasserquelle befand. Die Weitergabe dieser Informationen an nur bestimmte Personen der eigenen Gemeinschaft konnte möglicherweise das Überleben einer Gruppe sichern [101].

Dem deutschen Grundgesetz ist der Begriff der *Privatsphäre* als solcher nicht bekannt, allerdings wird das, „was man umgangssprachlich als ‚Privatsphäre' bezeichnen könnte, durch das Zusammenwirken mehrerer grundrechtlicher Normen geschützt, insbesondere durch das allgemeine Persönlichkeitsrecht aus Art. 2 Abs. 1 i.V. mit Art. 1 Abs. 1 Grundgesetz (GG), das Post- und Fernmeldegeheimnis aus Art. 10 GG und die Unverletzlichkeit der Wohnung aus Art. 13 GG" [97]. Privatsphäre ist ein vielschichtiges Phänomen und ein „Bedeutungsraum, in dem je nach System verschiedene Handlungen, Situationen, Zustände mentaler oder körperlicher Art des oder der Subjekte stattfinden, die in historisch und sozial variablem Ausmaß der Kontrolle des Außenraums entzogen werden" [98].

Rössler bricht die Komplexität des Privaten auf drei Dimensionen herunter: *informationelle Privatheit*, *dezisionale Privatheit* und *lokale Privatheit* [102]. Bei der informationellen Privatheit „geht es um Daten über eine Person, also generell darum, was andere über mich wissen". Private Entscheidungen und Handlungen fallen in den Bereich der dezisionalen Dimension der Privatheit „und steht die Privatheit meiner Wohnung zur Debatte, dann rede ich von lokaler Privatheit", schreibt Rössler [103]. „Trotz der Heterogenität der Verwendungsweisen des Privaten und trotz der Unterschiedlichkeit" dieser drei Dimensionen des Privaten tragen diese dazu bei, „einen gemeinsamen Nenner ausmachen", so Rössler [103]. Dieser besteht darin, die Frage „was diese Privatheit jeweils schützen soll" zu beantworten: „Privatheit schützt die individuelle Freiheit und Autonomie von Personen" [103].

Demnach lässt sich nun folgendes zur Privatsphäre festhalten: Privatsphäre umfasst einen Zustand des sich Zurückziehens mentaler oder auch physischer Art, was wiederum wichtig für Autonomie, Selbstentwicklung, Gesundheit und Wohlergehen einer Gesellschaft ist [98, 102, 104, 105, 106, 107]. In diesem gesellschaftlich akzeptierten und durch rechtliche Grundlagen gestützten Zustand, hat ein Individuum das Recht sich selbst (physisch) zurückzuziehen, seine Gedanken, Gefühle, Gewohnheiten oder Haltungen zurückzuhalten oder diese mit vom Individuum selbst bestimmten anderen Individuen teilweise oder ganz zu teilen.

Trotz der hier erfolgten Hintergrundinformationen besteht weiterhin die Notwendigkeit, genauer zu spezifizieren was Privatsphäre im Kontext von Informationssystemen meint und wie Software Engineers Privatsphäre operationalisieren können. Daher wird das hier herausgearbeitete Hintergrundwissen zur Privatsphäre später noch durch eine Definition des Begriffs Privatsphäre in Kapitel 6 ergänzt. Für den weiteren Verlauf der Arbeit sind die hier gegebenen Informationen aber vorerst ausreichend. Es sei an dieser Stelle noch ergänzt, das streng genommen unterschieden werden muss zwischen *Privatsphäre* und *Privatheit im digitalen Raum*, also der **Online-Privatsphäre**. Da diese Arbeit sich aber ausschließlich mit Konzepten und Maßnahmen der Online-Privatsphäre beschäftigt, sofern es nicht explizit anders angegeben wird, verwende ich aus Gründen der Lesbarkeit nur den Begriff der *Privatsphäre*, meine damit

aber die Online-Privatsphäre (informationelle Privatheit). Eine Abgrenzung beider Begriffe und eine Definition der Online-Privatsphäre, im Kontext des Software Engineerings, gebe ich zu gegebener Zeit in Abschnitt 6.1.2.

2.5.1 Privacy Attitudes und das Privacy-Paradoxon

Die individuelle Haltung/Einstellung zur Privatsphäre ist wichtig, „wenn es darum geht, das Verhalten der Menschen in Bezug auf die Privatsphäre zu verstehen“ [108]. Diese Haltung oder auch Einstellung wird als *Privacy Attitude* bezeichnet. Die Berücksichtigung von Privacy Attitudes spielt eine wichtige Rolle beim Design von digitalen Systemen, da aus den unterschiedlichen Privacy Attitudes der verschiedenen Endbenutzer auch unterschiedliches Verhalten in Bezug auf die eigene Privatsphäre im Umgang mit dem System resultiert. Daher sollte ein System, um seinen Nutzern entsprechenden Mehrwert zu bieten, auch ihre unterschiedlichen Privacy Attitudes mit einbeziehen [109]. Haltung und Verhalten in Bezug auf die eigene Privatsphäre sind nicht immer im Einklang und Nutzer verhalten sich häufig gegenteilig ihrer eigenen Haltung. In der Literatur ist diese Phänomen als *Privacy-Paradoxon* bekannt. Dieser Begriff wurde von Barnes [110] geprägt und ist von vielen anderen [111, 112, 113, 114, 115] gut erforscht. Kurz gesagt besagt das Privacy-Paradoxon: „Ich bin mir bewusst, dass meine Privatsphäre verletzt wird, dennoch nutze ich diesen Dienst weiterhin“. Das Privacy-Paradoxon betrifft alle Generationen [116] und „kann nicht allein auf ein mangelndes Verständnis oder ein mangelndes Interesse am Datenschutz zurückgeführt werden“ [113].

2.5.2 Privacy Awareness

Der Begriff *aware* bedeutet laut dem Cambridge Dictionary:[7] „knowing that something exists, or having knowledge or experience of a particular thing“, sich also einer Sache bewusst bzw. gewahr sein. Mit Blick auf die Online-Privatsphäre bedeutet es also, sich bewusst zu sein, dass persönliche Information in einem Interaktionskontext mit einem digitalen System und/oder Dritten anfallen und verarbeitet werden. Etwas formaler ausgedrückt und in Anlehnung an Pötzsch [114]: *Privacy Awareness* liegt vor, wenn ein Benutzer sich im klaren ist, dass seine persönlichen Informationen mit anderen (System(e) und/oder andere Nutzer) geteilt werden, wann das geschieht, welche Informationen das sind und in welchem Umfang diese erhoben und verarbeitet werden. „Vollständige“ Privacy Awareness kann aus meiner Sicht nur sehr begrenzt bzw. in selten Fällen vorliegen, da es nicht immer vollständig ersichtlich und transparent ist, ob ein System Daten sammelt, welche das sind oder wie diese genutzt werden.

[7] https://dictionary.cambridge.org/dictionary/english/

Man spricht in der Literatur aber auch schon von Privacy Awareness, wenn einem Benutzer klar ist, dass dieser in einem gegebenem Interaktionskontext persönliche Informationen preisgeben könnte [112, 117, 118, 119]. Ziel vieler privatsphärefördernden Mechanismen ist eine Sensibilisierung für bzw. Stärkung der Privacy Awareness des Endbenutzers. Ebendieses Ziel verfolgt auch mein hier vorgestelltes Konzept.

2.5.3 Privatsphäreaspekte

Das Cambridge Dictionary definiert einen Aspekt als „ein Teil einer Situation, eines Problems oder eines Themas, etc.“. Das Oxford Advanced American Dictionary[8] definiert einen Aspekt als „ein bestimmter Teil oder ein bestimmtes Merkmal einer Situation, einer Idee, eines Problems usw.; eine Art und Weise, in der es betrachtet werden kann“. Ein Aspekt ist also etwas, das einer Einheit (Subjekt, Objekt, Situation usw.) zugeordnet werden kann und einen Teil des Ganzen ausmacht. In diesem Sinne zählt grundsätzlich alles, was die Privatsphäre einer Person betrifft, als Aspekt der Privatsphäre. Das können zum Beispiel die Gedanken oder Gefühle, aber auch Gewohnheiten oder Haltungen einer Person sein. In Bezug auf die Online-Privatsphäre (Abschnitt 6.1.2) beziehen sich Privatsphäreaspekte zudem auch auf Daten und/oder Informationen von/über eine Person. Dies können neben Name, Adresse sowie Bankdaten, aber auch der Standort (GPS) einer Person sein, also Daten, die die Identifikation einer Person ermöglichen. Primär im US-amerikanischen Raum hat sich der Begriff der Personally Identifiable Information (PII) weitestgehend etabliert. Unter PII versteht man „alle Informationen über eine Person […], einschließlich (1) aller Informationen, die zur Unterscheidung oder Rückverfolgung der Identität einer Person verwendet werden können, wie z.B. Name, Sozialversicherungsnummer, Geburtsdatum und -ort, Mädchenname der Mutter oder biometrische Aufzeichnungen; und (2) alle anderen Informationen, die mit einer Person verknüpft sind oder verknüpft werden können, wie z. B. medizinische, Bildungs-, Finanz- und Beschäftigungsinformationen“ [120]. Mein Begriff des Privatsphäreaspekts umfasst ebenfalls diese Informationsmenge der PII, geht aber über den reinen Datenbezug hinaus, da auch Gedanken und Haltungen eingeschlossen sind.

2.6 Datenschutzgrundverordnung und Datenschutzerklärungen

Der Datenschutz in der Bundesrepublik Deutschland wird durch das Bundesdatenschutzgesetz geregelt. Es ergänzt und präzisiert die Datenschutzgrundverordnung (DSGVO) in den

[8] https://www.oxfordlearnersdictionaries.com

Bereichen, die den einzelnen Mitgliedstaaten der Europäischen Union (EU) überlassen sind. Entstanden ist das Datenschutzrecht „als Reaktion auf das Sammeln und Bearbeiten von Personendaten durch staatliche Behörden“ [121]. Als das weltweit erste formelle Datenschutzgesetz gilt das hessische Datenschutzgesetz von 1970 [121]. Die ersten Datenschutzgesetze regelten ausschließlich „das Bearbeiten von Personendaten durch Behörden“ und wurden auf das Verarbeiten von Personendaten durch Unternehmen 1977 im Bundesdatenschutzgesetz ausgedehnt [121].

2.6.1 Datenschutzgrundverordnung

Das aktuelle Datenschutzrecht in Deutschland und auch in der gesamten EU ist durch die DSGVO [122] geregelt. Die DSGVO ist geprägt durch den Grundsatz des Verbots der Datenverarbeitung ohne eine Einwilligung. Danach ist die Verarbeitung personenbezogener Daten grundsätzlich verboten, es sei denn, es liegt eine gesetzliche Erlaubnisgrundlage oder die Einwilligung der betroffenen Person vor. Damit soll sichergestellt werden, dass der Einzelne in der Regel selbst über die Preisgabe und Verwendung seiner personenbezogenen Daten entscheiden kann und somit sein Grundrecht auf den Schutz personenbezogener Daten (Art. 8 Charta der Grundrechte der EU [123]) wahrnehmen kann [124].

Um dieses Grundrecht zu schützen und durchzusetzen, verlangt das europäische Datenschutzrecht, dass die Nutzer umfassend über die Datenverarbeitung durch Software und andere Anwendungsprogramme informiert werden. Daher schreibt die DSGVO in unserer Gesetzgebung vor, dass die Datenverarbeitung *fair* und *transparent* erfolgen muss. Konkretisiert wird dies in den Artikeln 12 und 13 der DSGVO, die einerseits umfassende Informationen fordern, gleichzeitig aber auch eine *einfache Sprache* verlangen. Zu den obligatorischen Angaben gehören: die Kontaktdaten des für die Verarbeitung Verantwortlichen und eine Erläuterung der Rechte der betroffenen Person, die für die einzelnen Verarbeitungsvorgänge verwendeten Daten, die Begründung für die einzelnen Verarbeitungsvorgänge, der Zweck der Verarbeitung und die Dauer der Speicherung der Daten. Darüber hinaus können fakultative Informationen über die Weitergabe von Daten an Dritte, die Übermittlung in sogenannte Drittländer (außerhalb der EU) oder die Verwendung automatisierter Einzelentscheidungen gegeben werden. Diese Informationen werden jeweils durch die rechtliche Begründung ergänzt.

2.6.2 Fair Information Practices: Überblick weiterer Regularien

Der Vollständigkeit wegen sei an dieser Stelle noch erwähnt, dass es neben der DSGVO, die als umfassendste Datenschutzregelung gilt [125], weitere Empfehlungen oder Leitlinien zum Umgang mit persönlichen Daten gibt, die sogenannten *FIPPs*. Laut Schwartz [126] zählen die

FIPPs zu den Grundbausteinen des modernen Datenschutzrechts. Sie gründen auf dem *Privacy Act von 1974* [127], der sich zum Ziele setzte, „das Interesse des Einzelnen an der Privatsphäre zu wahren und gleichzeitig den legitimen Bedarf der Regierung an Informationen anzuerkennen“ [128]. Die FIPPs „sind eine Sammlung allgemein anerkannter Grundsätze, die von den Behörden bei der Bewertung von Informationssystemen, Prozessen, Programmen und Aktivitäten, die sich auf die Privatsphäre des Einzelnen auswirken, angewandt werden. Bei den FIPPs handelt es sich nicht um Vorschriften, sondern vielmehr um [allgemeine Grundsätze], die von jeder Behörde entsprechend ihrem jeweiligen Auftrag und den Anforderungen an das Datenschutzprogramm angewandt werden sollten“ [10]. Die FIPPs sind in Tabelle A.1 und Tabelle A.2 aufgelistet.

Die DSGVO ist bislang einzigartig, nicht zuletzt was die Regelung der Datenverarbeitung über Landesgrenzen hinweg betrifft. In den United States of America (USA) gibt es seit 2018 den California Consumer Privacy Act (CCPA) [129, 130]. Zielsetzung des CCPA ist, den Konsumenten mehr Kontrolle über ihre privaten Daten, die von Unternehmen bei ihren Geschäftsprozessen gesammelt und verarbeitet werden zu geben. Der CCPA verankert eher allgemein das Recht der Konsumenten bei Verarbeitung ihrer Daten durch Unternehmen und basiert nicht auf den FIPPs. Mit dem California Privacy Rights Act (CPRA) erhält der CCPA eine Art Verfassungszusatz [131]. Der CPRA ist eine Art Pendant zur DSGVO und regelt die Rechte der Benutzer im Internet in Bezug auf ihre privaten Daten und deren Verarbeitung. Der CPRA basiert auf den FIPPs und weißt somit einige Ähnlichkeiten zur DSGVO auf. Sowohl CCPA als auch CPRA sind Regelungen des Bundesstaats Kalifornieren und somit in den USA nicht landesweit gültig.

Die International Standard Organization (ISO) hat mit dem ISO 29100 Standard [132] ein High-level Framework geschaffen, dass den Schutz personenbezogener Daten, basierend auf der Umsetzung von Datenschutzprinzipien spezifiziert. „Die in diesem Standard beschriebenen Datenschutzprinzipien wurden von bestehenden Prinzipien abgeleitet, die von einer Reihe von Staaten, Ländern und internationalen Organisationen entwickelt wurden“ [131].

2.6.3 Datenschutzerklärungen

Eine Datenschutzerklärung (DSE) beschreibt, „welche Personendaten und sonstigen Informationen von einer Organisation auf welche Weise und zu welchem Zweck bearbeitet werden. Sie dient der Transparenz und Aufklärung der betroffenen Personen und ermöglicht Ihnen die Ausübung ihrer datenschutzrechtlichen Rechte“ [133]. Letzteres wird als das Prinzip von *Notice und Choice* bezeichnet [134]. Lepperhoff und Petersdorf [135] ergänzen zudem, das

Unternehmen mit einer DSE signalisieren, „wozu sie persönliche Angaben nutzen. Die so herbeigeführte Transparenz stellt eine wichtige Grundlage für Vertrauen dar. Aus diesem Grund liegen Datenschutzerklärungen im Eigeninteresse von Unternehmen“.

DSEs wie sie aktuell gestaltet sind, stehen immer noch im Kontrast zu dem, was die DSGVO einfordert [136, 137, 138] und es existieren bislang fast keine anderweitig umgesetzten Hinweise oder Informationen für den Endnutzer.

2.6.3.1 Tracking-Techniken und Cookie Banner

Ergänzend zu den DSEs gibt es noch eine weitere Art, die Nutzer über Datenpraktiken zu informieren, wenn auch nur rudimentär, da hier lediglich ein Hinweis auf eine mögliche Datenverarbeitung gegeben wird, Details dazu aber i.d.R. in den DSEs zu finden sind. In Kenntnis gesetzt werden die Nutzer über sogenannte *Cookie Banner* oder auch synonym *Cookie Consent Notices*. Im Wesentlichen wird zwischen funktionalen Cookies (auch technisch notwendig genannten Cookies) und nichtfunktionalen Cookies unterschieden [137]. Hierbei werden die funktionalen Cookies benötigt, um zum Beispiel das korrekte Darstellen einer Internetseite zu gewährleisten, wo hingegen die nichtfunktionalen Cookies einen optionalen Wert haben und das Ablehnen dieser keinen Einfluss auf das Benutzererlebnis einer Internetseite oder eines digitalen Dienstes hat. Die nichtfunktionalen Cookies werden häufig von Drittanbietern zu Werbezwecken und/oder zur Erstellung von Surf-Profilen eingesetzt. Eine Internetseite muss seit 2009 [139] Benutzer über die eigenen Cookie-Richtlinien aufklären. Dies geschieht über Cookie Banner. Benutzer können hier der Verwendung von Cookies zustimmen und ggf. nichtfunktionale Cookies ablehnen. Die Implementationen dieser Cookie Banner weisen allerdings eklatante Schwächen auf, scheinen bewusst irreführend gestaltet zu sein und nutzen zudem Aspekte des Privacy Paradoxons aus, um Nutzer schnell zur Zustimmung der Verwendung von nichtfunktionalen Cookies zu bewegen [38, 137, 140, 141, 142].

2.7 Vertrauen & Vertrauenswürdigkeit

Ohne ein Mindestmaß an Vertrauen würden wir unseren Alltag kaum bewältigen können. Unbestimmte Ängste und ein damit einhergehendes Gefühl der Angst würden uns lähmen [143]. Vertrauen ist zentral für zwischenmenschliche Beziehungen; diese wären ohne Vertrauen nicht denkbar. Überall dort, wo Interdependenz, Ungewissheit und Risiko herrschen, übernimmt Vertrauen die Rolle eines Vermittlers [144]. Vertrauen ist daher aus gesellschaftlichen und auch wirtschaftlichen Gründen entscheidend [145]. Es ist wichtig für das gesellschaftliche Zusammenleben sowie dessen Solidarität und ist ein entscheidender Faktor für die Leistungsfähigkeit und Entwicklung von Wirtschaftssystemen [146].

2.7.1 Abgrenzung von Vertrauen & Vertrauenswürdigkeit im Kontext des Software Engineering

Beim RE und dem Design von Informationssystemen spielt Vertrauen ebenfalls eine wesentliche Rolle [147, 148, 149]. Viele sehen in dem Konzept der Erklärbarkeit ein geeignetes Mittel, um das Vertrauen in ein System bzw. das Vertrauen der Stakeholder zu stärken [150, 151, 152, 153]. Vor diesem Hintergrund könnte es sinnvoller sein, Anforderungen an die Erklärbarkeit zu erheben, als direkt Anforderungen an das Vertrauen zu stellen.

Wichtig ist hierbei aber zu differenzieren. Auch wenn Erklärungen dazu beitragen können, das Vertrauen in ein System zu erhöhen, bedeutet dies nicht notwendigerweise, dass ein System dann auch vertrauenswürdig (englisch: *trustworthy*) ist. Dies hängt wiederum sowohl von der Motivation als auch von der Umsetzung ab. Gegeben sei beispielsweise ein System, das Gesetzgeber und Juristen bei der Entscheidungsfindung rechtlicher Angelegenheiten unterstützt. Hier ist es wichtig, dass das System verlässliche, faire und ethische Empfehlungen liefert, die innerhalb der Rechtsprechung vertrauenswürdig sind. Erklärungen können hier nützlich sein, um diese Ziele zu unterstützen und den Argumentationsprozess des Systems - das *Reasoning* - verständlich und nachvollziehbar zu machen, aber zusätzliches gerechtfertigtes Vertrauen (englisch: *warranted trust*) in die Interna des Systems ist notwendig, um die Vertrauenswürdigkeit des Systems zu gewährleisten. Vertrauenswürdigkeit kann als Mediator für Vertrauen fungieren [154]. Das Beispiel verdeutlicht, dass es bei der Entwicklung von Systemen durchaus wichtig sein kann, zwischen Vertrauen und Vertrauenswürdigkeit zu unterscheiden.

Nach Kästner et al. ist [154] *Vertrauen* „eine Haltung, die ein Stakeholder gegenüber einem System hat“. Im Gegensatz dazu beschreiben die Autoren *Vertrauenswürdigkeit* als „Eigenschaft eines Systems: intuitiv ist ein System für einen Stakeholder vertrauenswürdig, wenn es für den Stakeholder gerechtfertigt ist, Vertrauen in das System zu setzen“. Vor diesem Hintergrund sollte ein System in einem bestimmten Kontext richtig funktionieren. Für das obige Beispiel bedeutet dies, dass das System, das Gesetzgeber und Juristen bei der Entscheidungsfindung unterstützt, fair und gerecht sein muss und niemanden diskriminieren darf und absolut gesetzeskonform agiert. Dies *muss* gewährleistet sein. Dann wäre das System vertrauenswürdig. Insbesondere im Hinblick auf den Schutz der Privatsphäre ist es wichtig, zwischen Vertrauen und Vertrauenswürdigkeit zu unterscheiden, denn wenn die Endnutzer Systemen vertrauen sollen, müssen sie auch sicher sein, dass diese Systeme vertrauenswürdig sind.

2.8 Verwandte Arbeiten

Die in diesem Kapitel bisher genannten Arbeiten stehen alle in Bezug zu meiner Dissertation und ich erachte diese daher als verwandte Arbeiten, ebenso wie die zahlreichen Arbeiten in den nachfolgenden Kapiteln, auf denen sich meine Forschungen begründen. Darüber hinaus möchte ich hier dennoch eine kleine Auswahl verwandter Arbeiten vorstellen, die komprimiert die Bereiche der Erklärbarkeit und der Datenschutzerklärungen (DSEs) betrachten, um dem Leser einen Überblick über die Forschungen in diesen Bereichen zu verschaffen.

2.8.1 Erklärbarkeit

Erklärbarkeit ist im Bereich des maschinellen Lernens (ML) weder von einer einheitlichen noch konsistenten Terminologie umgeben [155]. Es koexistieren diverse, teils synonym verwendete Begriffe wie *interpretability* [156, 157], *scrutability* [158] oder auch *transparency* [159]. Die Erklärbarkeit ist besonders in der XAI-Domäne [160] ein intensiv beforschtes Thema. Dabei geht es im wesentlichen darum, dass „Innenleben“ eines Systems besser zugänglich und die Ergebnisse wie Vorhersagen oder Empfehlungen, berechenbar bzw. bewertbar (englisch: *assessable*) zu machen. Es geht also darum, die Blackboxness[9] eines Systems transparenter und verständlicher zu gestalten. Das Ziel dabei sind Qualitätsziele wie *accountability* [162], *fairness* [163], *trust* [164], *understandability* [153] etc. zu erreichen. Erklärbarkeit, als eigenständige NFR [165, 52], steht in Wechselwirkung mit vielen anderen Qualitätsaspekten und kann diese sowohl positiv als auch negativ beeinflussen [4]. Zum Beispiel können zu viele oder schlecht designte Erklärungen einen negativen Einfluss auf die User Experience eines Systems haben [166, 167]. Ebenso kann zu viel Transparenz das Vertrauen in ein System sowie dessen Verständlichkeit schwächen [168, 169]. Daher ist es wichtig, beim Design erklärbarer Systeme die Adressaten der zu gebenden Erklärungen zu berücksichtigen, sowie deren unterschiedlichen Bedarfe.

Langer et al. [170] stellen ein Modell vor, was die verschiedenen Stakeholder von erklärbaren Systemen klassifiziert und ihre unterschiedlichen Desiderata (NFRs) in Beziehung setzt. Das Modell soll bei der Bewertung, Anpassung, Auswahl und Entwicklung von Erklärungsansätzen helfen, die Wünsche und Bedarfe der unterschiedlichen Stakeholder in Einklang zu bringen und zu erfüllen. Chazette et al. [171, 172] machen ebenfalls Vorschläge die Entwicklung erklärbarer Systeme zu verbessern. Hierzu stellen die Autoren u.a. ein Framework vor, dass den Entwicklungsprozess von der Erhebung der Erklärbarkeitsanforderungen über deren Implementation bis hin zum Testen unterstützen soll.

[9] Der Begriff *Blackboxness* bezieht sich darauf, dass einige ML-Modelle für uns so geheimnisvoll wie eine Blackbox sein können, da ihr Innenleben undurchsichtig und schwer zu interpretieren und zu erklären ist [161].

Einfach nur Erklärungen in ein System zu implementieren ist nicht zielführend und kann das System auf verschiedene Arten schwächen und schädigen, wie bereits erwähnt. „Die Bewertung der Erklärbarkeit erlaubt eine Beurteilung der Qualität von Erklärungen und ermöglicht den Vergleich verschiedener Erklärungsvarianten“ [173]. Die Evaluierung der Eignung von Erklärungen ist nicht trivial, da es hierbei u.a. davon abhängt welche Qualitätsziele durch Erklärungen erreicht werden sollen. Deters et al. [173] setzen bei der Bewertung der Erklärbarkeit im Hinblick auf vordefinierte Ziele das Konzept zielorientierter Heuristiken ein, um zu prüfen, ob das jeweilige Ziel mittels der Erklärbarkeit erreicht wurde.

2.8.2 Datenschutzerklärungen und -hinweise

Untersuchungen haben gezeigt, dass die durchschnittliche DSE zwischen 2600 und 5200 Wörter umfasst, die oft in komplizierter Sprache verfasst, unverständlich und uneindeutig formuliert sind [174, 48, 175]. Dennoch sind sie nach wie vor das wichtigste Mittel zur Information und Aufklärung der Nutzer über die Datenverarbeitung [48]. Bereits 2002 wurden mit dem freiwilligen Standard *Platform for Privacy Preferences* (P3P) erste Anstrengungen unternommen, um DSEs maschinenlesbar zu machen [176]. Ziel war es, die DSEs für die Nutzer leichter zugänglich zu machen. Ein Ansatz, der diesen Standard nutzte, wurde als *P3P-Benutzeragent* [177] vorgestellt, der in der Lage war, DSEs auf der Grundlage dieses Standards zu verarbeiten. Der Standard wurde jedoch nie eingeführt.

Es gibt verschiedene Ansätze und Tools DSEs automatisiert aufzubereiten bzw. zu analysieren. *Privee* [178] ist das erste mir bekannte Tool zur automatischen Analyse von DSEs. Privee nutzt dafür maschinelles Lernen sowie das Crowd-Sourcing-Datenschutzanalyse-Framework ToS;DR[10]. Harkous et al. [179] stellen ein Framework (Polisis) zur automatisierten Analyse von DSEs vor. *PriBot*, ein ChatBot-ähnliches System baut auf Polisis auf. Es ermöglicht dem Nutzer, aktiv Fragen zu einer DSE zu stellen, wie z.B. „Werden meine Daten an Dritte weitergegeben?“. *Mobile App Privacy System* – kurz MAPS – [180] ist eine Anwendung, um Konformität zwischen Code und den entsprechenden DSEs mobiler Android Apps zu prüfen. Die Autoren deckten mit Hilfe von MAPS „zahlreiche Hinweise auf mögliche Verstöße“ [180] gegen die DSEs auf. *PolicyLint* [181] ist ein Tool, das potenzielle Widersprüche innerhalb von DSEs aufdecken kann. PrivacyGuide [182] ist ein nicht öffentliches Tool, ähnlich zu *PrivacyCheck v2* [183], das mit Hilfe von maschinellem Lernen und natürlicher Sprachverarbeitung DSEs in Anlehnung an die DSGVO analysiert und verarbeitet.

Neben der automatisierten Analyse von DSEs gibt es zahlreiche Arbeiten, die darauf abzielen, DSEs benutzerfreundlicher und nutzerorientierter zu gestalten [175, 184, 185, 186] und sie sogar kontextabhängig darzustellen [187, 188], d. h. sie so anzupassen, dass nur die für den

[10] https://tosdr.org/

jeweiligen Nutzungskontext relevanten Informationen angezeigt werden [189, 190]. Kulyk et al. [191] zeigten, dass die Bereitstellung geeigneter (einfacher und verständlicher) Informationen für die Nutzer, diese in die Lage versetzt, eine explizite Zustimmung (informed consent) zu geben und bessere Entscheidungen in Bezug auf ihre Privatsphäre zu treffen. In ihrer Arbeit verwendeten sie dazu einen Informations-Flyer als Medium, um die relevanten Datenschutzinformationen zu vermitteln. Shulman et al. [192] führten eine Studie durch, um herauszufinden wie und ob sich das Verhalten der Teilnehmer beim Offenlegen privater Informationen ändern, wenn den Nutzern Privatsphärehinweise gezeigt werden. Shulman et al. fanden heraus, „dass die Effektivität der Informierung der Nutzer durch Benachrichtigungen vom Zeitpunkt, Inhalt und Layout dieser Benachrichtigungen abhängen kann“ [192]. Ein weiterer interessanter Effekt der sich in den Ergebnissen zeigte – ähnlich wie beim Konsumverhalten wo z.B. Rabatte oder Treueprogramme [34, 35] einen Einfluss auf die Entscheidung der eigenen Privatsphäre haben können –, dass Neugier ein ähnliches Verhalten triggert. Ein hohes Maß an Neugier ließ Nutzer eher Informationen offenlegen, „während Rationalität mit mehr Rücksicht auf die Privatsphäre verbunden ist“ [192]. Das gleiche scheint für positive und negative Emotionen zu gelten. „Ein positiver Affekt wird mit geringeren Erwägungen zum Schutz der Privatsphäre in Verbindung gebracht, während ein negativer Affekt mit verstärkten Erwägungen zum Schutz der Privatsphäre einhergeht“, so Shulman et al. [192]. Kelley et al. [193] zeigten in ihren Studien, das eine Standardisierung in der visuellen Repräsentation von DSEs einen positiven Effekt auf das schnelle und treffsichere Auffinden spezifischer Informationen haben kann. Zum Einsatz kamen hier Kennzeichnungen in Anlehnung an Produktkennzeichnungen wie man sie von Lebensmitteln oder Elektrogeräten kennt [194]. Andere Ansätze versuchen Konsumentenvertrauen über Privatsphäresiegel (englisch: *privacy seals*) zu etablieren [195, 196]. Die Siegel werden dabei nach Prüfung durch Dritte vergeben. Ein Hindernis besteht hierbei allerdings darin, dass Nutzer weder wissen, was ein Anbieter tun muss, um ein solches Siegel zu erhalten, noch wie ein echtes Siege ausschaut [197]. Eine Herausforderung, sowohl bei den Kennzeichnungen als auch den Siegeln, bleibt allerdings für die Endbenutzer weiterhin bestehen, nämlich Details zu den Datenpraktiken der Anbieter eigenständig herauszufinden. Die Privatsphäresiegel und Kennzeichnungen können eine sinnvolle Ergänzung darstellen, um beispielsweise die Vertrauenswürdigkeit oder Datenkonsum eines Anbieter auf einen Blick zu erfassen. Als detaillierte und aufklärende Informationsquelle für Endbenutzer sind sie aber eher ungeeignet.

Der Vollständigkeit halber sei noch erwähnt, dass DSEs neben der Informierung über Datenpraktiken ebenfalls für das Requirements Engineering eine wichtige Rolle spielen [198]. Sie können als Quellen für Anforderungen dienen [199, 200]. Nicht verwunderlich ist, dass bei der automatisierten Extraktion von Anforderungen zu ähnlichen Probleme kommt, wie sie der Endnutzer hat: Mehrdeutigkeiten, die unterschiedliche Interpretationen zulassen [201, 202].

3

Konzeptualisierung des wissenschaftlichen Vorgehens

Dieses Kapitel beleuchtet den für diese Arbeit gewählten wissenschaftlichen Ansatz zur Beantwortung der Forschungsfragen und somit zur Erreichung des Forschungsziels. Der gewählte Ansatz orientiert sich an der Design Science Research (DSR) [57, 203] und der Information System Research (ISR) [2, 204, 205]. ISR ist eine angewandte Forschungsdisziplin, da hier „Theorien aus anderen Disziplinen wie Wirtschaft, Informatik und Sozialwissenschaften“ [206] zum Einsatz kommen, „um Probleme an der Schnittstelle von Informationstechnologie und Organisationen zu lösen“ [206]. Das Ziel der ISR besteht darin, „Wissen zu schaffen, das die Anwendung der Informationstechnologie für Management- und Organisationszwecke ermöglicht“ [2], um den zuvor genannten Problemen zu begegnen. Hierfür sollte die ISR laut Hevner und March [2] „Kreativität und Präzision der Designwissenschaft (Design-Science) mit der Empirie und Disziplin der Verhaltenswissenschaft verbinden.“

Das Paradigma der Designwissenschaft hat seinen Ursprung laut Hevner et al. [57] im Ingenieurwesen und in den *Wissenschaften des Künstlichen* (The Sciences of the Artificial, Simon 1996 [207]). Hierbei ist die Zielsetzung der Designwissenschaft, Artefakte oder auch Innovationen zu kreieren. Diese wiederum verkörpern die Ideen, technischen Fähigkeiten, Praktiken

sowie Produkte, die notwendig sind, um (a) die Analyse, (b) den Entwurf, (c) die Implementierung sowie (d) die Nutzung, von Informationssystemen auf eine effiziente Art und Weise durchzuführen [2, 57].

Im Kern bedeutet dies, dass der Beitrag der Forschung in der Domäne der Informationssysteme darin besteht, Theorien und Artefakte zu entwickeln, die zum einen auf realen Herausforderungen und Bedürfnissen beruhen (somit also relevant sind), zum anderen existierende wissenschaftliche sowie fachliche Erkenntnisse mit einbeziehen, um stringent, schlüssig und exakt zu sein.

3.1 Details zum Forschungsvorgehen

Abbildung 3.1 veranschaulicht das auf ISR und DSR gründende Forschgunsvorgehen, eingebettet in das ISR Framework. Die realen Bedürfnisse und Herausforderungen wurden anhand von vorhandener Literatur sowie unterstützend durch Umfragen identifiziert. Die vorhandene Wissensbasis verhalf dazu, Konzepte und gängige Praktiken der Softwarequalität, Erklärbarkeit und Privatsphäre zu ermitteln.

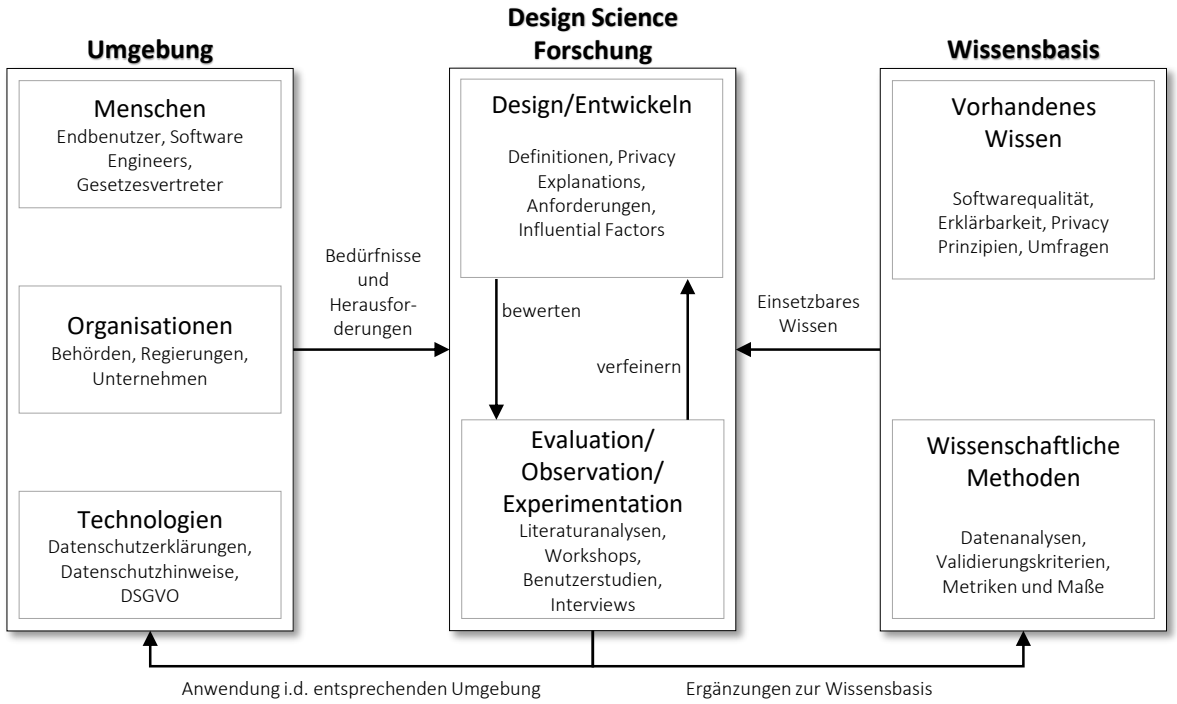


Abbildung 3.1: Information Systems Research Framework, in Anlehnung an [2]

Bei der Durchführung der Studien wurden multimethodische Ansätzen verfolgt, sowie Daten-Triangulation eingesetzt. Bezugnehmend auf Wilson [208], bedeutet der Begriff der Triangulation im Kontext von Forschungsmethoden, dass „dass bei der Forschung mehr als ein bestimmter Ansatz verwendet wird, um reichhaltigere, umfassendere Daten zu erhalten und/oder die Ergebnisse der Forschung zu bestätigen“ [208]. Der Begriff der Daten-Triangulation wurde von Flick [209] geprägt und gemeint ist u.a. das Verwenden unterschiedlicher Quellen von Daten mit dem Ziel „den Umfang, die Tiefe und die Kohärenz des methodischen Vorgehens“ [209, S.457] zu erhöhen und somit höhere Zuverlässigkeit der Ergebnisse zu erreichen. Triangulation kann sowohl für quantitative als auch qualitative Forschung genutzt werden und Wilson gibt zu bedenken, dass es so scheint, als sei es ein anderer Term für *mixed-methods*-Ansätze [208]. Daher werden im Rahmen dieser Arbeit die Begriffe multimethodischer Ansatz und mixed-method-Ansatz synonym verwendet. Die genauen Ansätze, die bei den jeweiligen Studien zum Einsatz kamen, werden jeweils in den entsprechenden Kapiteln vorgestellt und erläutert.

3.2 Einbettung des Vorgehens in das ISR Framework

Literatur und Umfragen im Bereich der Erklärbarkeit und Privatsphäre haben dazu beigetragen, das relevante Problem zu identifizieren.

> ***Relevantes Problem*** Informationen über die Verwendung persönlicher Daten erhalten Benutzer digitaler Systeme aktuell primär über DSEs. Bezugnehmend auf die Problemstellung (siehe 1.1) dieser Dissertation, verfehlen diese aber im Hinblick auf Informierung und Aufklärung der Nutzer ihren Zweck [35, 43, 44, 45, 46, 47, 48, 45]. Organisationen auf der anderen Seite könnten ebenfalls davon profitieren, einfachere und nachvollziehbarere Hinweise bezüglich ihrer Datenpraktiken bereitzustellen. Denn Nutzer sind durchaus beispielsweise durchaus gewillt, von Vorteilen im Tausch für Ihre privaten Daten zu profitieren [33, 34, 36]. Allerdings möchten sie auch darüber dementsprechend informiert werden [50, 51]. Nachvollziehbare Aufklärung über Datenpraktiken ist also ein relevantes IT-Problem sowohl von der Perspektive der Unternehmen als auch der Endbenutzer.

Basierend auf dem festgestelltem (relevanten) Problem werden Lösungen entwickelt – die Theorien und Artefakte. Diese bauen auf dem festgestelltem Problem auf. Die Theorien im Rahmen dieser Dissertation sind von empirischer Natur. Bezugnehmend auf Shadish et al. [210] setzt eine empirisch basierte Theorie, wie der Name schon vermuten lässt, empirische Mittel ein. Dadurch ist es möglich, analytisch zu verallgemeinern, um somit Prozesse oder Phänomene zu erklären, in denen eine statistische Verallgemeinerung nicht wünschenswert oder

möglich ist. Das kann z.B. bei Fallstudien über Populationen hinweg der Fall sein oder auch bei Experimenten aus den Bereichen der Sozial- und Verhaltenswissenschaften mit denen sich das empirische Software Engineering wesentliche Eigenschaften teilt [211]. Laut Sjøberg et al. [211] kann eine Theorie sowohl für die Forschung als auch für die Praxis hilfreich sein. Eine Theorie hat im Grunde genommen auch immer etwas Praktisches, da sie Beobachtungen oder Phänomene voneinander differenziert, diese greifbar macht oder deutet. Während meiner Forschungen und den damit verbundenen Recherchen zur Beantwortung der Forschungsfragen (siehe Abschnitt 1.2), habe ich drei Theorien (T_n) entwickelt und sieben Artefakte (A_n) geschaffen.

Geschaffene Theorien und Artefakte Die erste Theorie T_1 – *Erklärbarkeit im Kontext der Privatsphäre* – ist ein Mix aus Analyse- und Erklärungstheorie [211], die zum einen das Konzept der Erklärbarkeit beschreibt sowie in einer Definition für erklärbare Systeme konzeptualisiert (A_1, Abschnitt 4.2) und zum anderen ein Modell (A_2, Abschnitt 4.3), das zur Beantwortung von **RQ1** führt. A_2 beschreibt die Beziehung zwischen Erklärbarkeit und Privatsphäre sowie im Einfluss stehender weiterer Qualitäten und deren Auswirkungen. Meine zweite entwickelte Theorie T_2 – *Online-Privatsphäre im Kontext des Software Engineerings* –, ist ebenfalls wie T_1 eine hybride Theorie. Sie beschreibt den Begriff der Online-Privatsphäre (A_3, Abschnitt 6.1.2) im Kontext des Software Engineerings und zeigt eine Möglichkeit auf, wie das Konzept der Erklärbarkeit eingespannt werden kann, um Privatsphäreaspekte auf nutzerzentrierte Art und Weise zu erklären (A_4, Abschnitt 6.1.3), was zur Beantwortung von **RQ2** führt. Dafür verbindet T_2 das konzeptionelle Modell A_2 das die verschiedenen Einflussfaktoren beschreibt, die bei der Erklärung von Privatsphäreaspekten von Relevanz sind und A_1. T_3 ist eine sogenannte *Design und Action*-Theorie [211] – *Erklärbarkeit und Privatsphäre in der Praxis*. T_3 umfasst nicht nur theoretische Konzepte, sondern liefert wichtige Artefakte, wie sich so genannte *Privacy Explanations*[1] in der Praxis umsetzen lassen und was dabei zu beachten ist. Hierfür stelle ich neben Anforderungen an Privacy Explanations (A_5, Abschnitt 7.2.3), ein Design-Konzept, bestehend aus dem Prinzip von Context, Content und Consent und dem Layered-Ansatz, als sechstes Artefakt (A_6, Abschnitt 7.2.2) vor und zeige mit dem *Trialog of Privacy* (A_7, Abschnitt 8.2.3) eine Möglichkeit auf, privatsphärebewusste Entscheidung im Entwicklungsprozess zu verankern. A_5 und A_6 führen schließlich zur Beantwortung von **RQ3**.

[1] Der Begriff „Privacy Explanation“ wird in Kapitel 6 definiert.

Die Theorien und Artefakte formen die Lösungen zum relevanten Problem und tragen zur Beantwortung meiner Forschungsfragen bei. Während der Evaluation wird geprüft, ob Theorien und Artefakte auch hilfreich dazu beitragen, reale Bedürfnisse und Herausforderungen zu befriedigen bzw. zu lösen.

> ***Evaluation und Verfeinerung*** Die Theorien und Artefakte wurden mit Hilfe von Literaturrecherchen, Workshops, Umfragen und Benutzerstudien entwickelt, bewertet und validiert. Die Ergebnisse der Evaluierungen wurden wiederum dazu verwendet, die Artefakte zu verfeinern und die Wissensbasis zu erweitern.

4

Das Konzept der Erklärbarkeit

Erklärbarkeit als solches wird schon seit langer Zeit von anderen Disziplinen wie der Philosophie oder Psychologie erforscht. Aber auch für die Informatik als Wissenschaft erschien bereits 1984 die erste große Arbeit hierzu [212]. Im Bereich der XAI, ein Teilgebiet der Informatik, ist Erklärbarkeit ebenfalls ein schon seit längerem sehr prominentes Forschungsthema. Denn besonders im Bereich des maschinellen Lernens sind Schlussfolgerungen sowie Entscheidungen dieser Systeme und der zugrunde liegenden Modelle häufig äußerst undurchsichtig und wirken daher wie mysteriöse Black-Boxen [213, 161]. Diese sogenannte *Blackboxness* gepaart mit der gesellschaftlich engen Verflechtung digitaler Systeme, regte einen Anstoß für eine Diskussion über Ethik und Transparenz dieser Systeme an [214]. Verantwortungsvolle Datenpraktiken, Privatsphäre und Sicherheit sind hierbei nur einige neben vielen weiteren Bedenken. Daher ist es wichtig zu verstehen, wie diese Bedenken beim Design und Entwurf von Software-Systemen zu berücksichtigen sind.

Erklärbarkeit als NFR wird zunehmend als Hilfsmittel hin zu einer Lösung gesehen, dieser „Blackboxness“ zu begegnen und dem Mangel an Transparenz eines Systems entgegenzuwirken [165]. Erklärbarkeit kann Gefühle der Frustration vermeiden [215] und wirkt sich zudem auf das Vertrauen in ein System aus [216]. Darüber hinaus haben andere Studien gezeigt, dass Erklärbarkeit auch mit anderen Qualitätsaspekten wie Überprüfbarkeit (englisch: *auditability*) und Bedienbarkeit verbunden ist [165, 217, 218]. Ebenso wie bei anderen NFRs auch, ist

Erklärbarkeit als Qualität schwer zu elizitieren und zu validieren. Das liegt an der interaktiven, relativen und subjektiven Natur von NFRs, was für Requirements Engineers oft eine große Herausforderung ist [219]. Das hat seinen Ursprung darin, dass die Interpretation von NFRs von einer Reihe von Faktoren abhängt, „wie z. B. dem jeweiligen System, das entwickelt wird“ [219] und der Art und Weise der Beteiligung der Stakeholder. „NFRs können von verschiedenen Personen und in verschiedenen Kontexten, in denen das System entwickelt wird, unterschiedlich betrachtet, interpretiert und bewertet werden. Folglich sind die positiven oder negativen Beziehungen zwischen ihnen nicht immer offensichtlich.“ [219]. Darüber hinaus sind NFRs dafür bekannt, „einen *make-or-break*-Status im Software-Entwicklungsprozess zu haben, lassen sich aber nur schwer formell behandeln“ [220]. Da Qualitätsaspekte in der realen Welt verwurzelt und somit von abstrakter Natur sind, ist es unerlässlich, dass sie genau verstanden werden, um in korrekt spezifizierte Anforderungen übersetzt werden zu können. Dafür ist aber wichtig für den Erfolg im Software-Entwicklungsprozess, dass ein gemeinsames Verständnis aller Stakeholder gerade in Bezug auf NFRs vorhanden ist [221]. Sowohl in agilen als auch in traditionellen Entwicklungsprojekten wird das gemeinsame Verständnis durch die Erhebung von Anforderungen und anschließender Formalisierung in einer Spezifikation, Stories, Upfront-Testfällen oder anderen Design-Artefakten erreicht [79]. Eine Möglichkeit für ein gemeinsames Verständnis unter und mit allen Beteiligten bei der Software-Entwicklung zu sorgen besteht darin, sicherzustellen, dass alle die verwendeten Begriffe und Fachtermini auf die gleiche Art und Weise verstehen bzw. wahrnehmen [222, 223].

Da Erklärbarkeit eine neu entstandene NFR ist, besteht noch kein ausreichend strukturiertes Wissen über diese Qualität. Das Ziel dieses Kapitels ist, diese Lücke mit Wissen um eine gemeinsame Terminologie und Semantik zu schließen, um die Diskussion und Analyse der Erklärbarkeit während des RE-Prozesses zu erleichtern. Darauf aufbauend wird dann die Beziehung von Erklärbarkeit und Privacy eingehender untersucht. Dafür beginnt das Kapitel mit der Beschreibung unseres[1] Forschungsvorgehens (Abschnitt 4.1). In Abschnitt 4.2 wird eine Definition für erklärbare Systeme vorgestellt und das Kapitel schließt mit der Betrachtung der Beziehung zwischen Erklärbarkeit und Privacy in Abschnitt 4.3.

Zugehörige Publikationen Der in diesem Kapitel vorgestellte Forschungsbeitrag entstand in Kollaboration mit zwei anderen Forschern: Larissa Chazette und Timo Speith. Dieses Kapitel gründet auf der gemeinsamen Zusammenarbeit und die Ergebnisse dazu wurden in [4] und [3] veröffentlicht, worauf dieses Kapitel basiert.

[1] Aufgrund der Kollaboration mit meinen Forschungskollegen verwende ich das „wir“ in diesem und in anderen Kapiteln dieser Dissertation, wenn die Forschung durch Kollaboration entstand.

4.1 Forschungsvorgehen

Das hier angewandte Forschungsvorgehen besteht aus zwei einander ergänzenden Vorgehen, die zu der Definition für erklärbare Systeme und dem Modell der Beziehungen von Erklärbarkeit und Privatsphäre führten. Beide Forschungsvorgehen verfolgten einen multimethodischen Ansatz bestehend aus einer systematischen Datenerhebung und einer qualitativen Analyse der Daten. Nachfolgend erläutere ich zunächst das Vorgehen aus der Forschungskooperation mit Larissa Chazette und Timo Speith. Anschließend gehe ich ausführlich auf den Teil der eigenen Forschung ein, die zum Modell der Beziehungen von Erklärbarkeit und Privatsphäre führte.

4.1.1 Forschungsvorgehen - Definition erklärbare Systeme und Einfluss der Erklärbarkeit

Im Folgenden beschreibe ich unser Vorgehen in Grundzügen, welches in Abbildung 4.1 dargestellt ist. Weitere Details wie genaue Anzahl an selektierten Publikationen der SLR etc. sind in Anhang B zu finden. Zudem ist dieser Teil der Forschungskooperation bereits ausführlich in der Dissertation von Chazette [56] beschrieben.

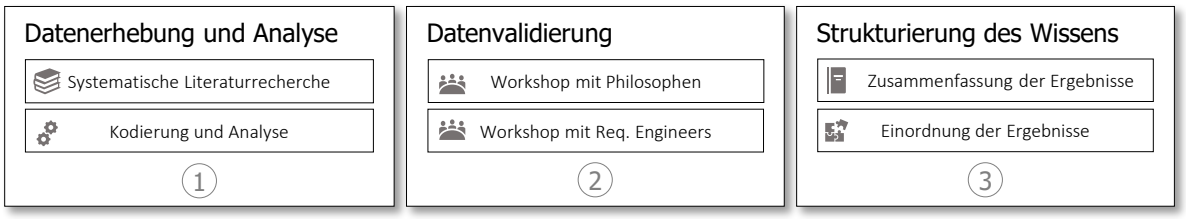


Abbildung 4.1: Übersicht des Forschungsvorgehens mit Bezug auf [3, 4]

4.1.1.1 Datenerhebung und Analyse

Systematische Literaturrecherche Wir haben uns für eine interdisziplinäre SLR entschieden, weil andere Disziplinen wie Philosophie und Psychologie bereits über eine jahrzehntelange Erfahrung im Bereich der Erklärbarkeit verfügen. Diesen Erfahrungsschatz anderer Disziplinen wollten wir nutzen, um zu schauen was das SE bzw. RE von diesen Disziplinen lernen kann. Bei der Durchführung der SLR folgten wir den Richtlinien von Kitchenham et al. [224] und Wohlin [225]. Die Suchstrategie der SLR beinhaltete eine manuelle Suche, gefolgt von einem Snowballing-Prozess. Bei der manuellen Suche inspiziert ein Forscher spezifische (relevante) Quellen wie Konferenzbände, Workshops oder Journals und zwar Ausgabe für Ausgabe [226]. Dafür haben wir zuerst relevante Quellen aus dem Bereich der Informatik,

Psychologie und Philosophie identifiziert und diese haben wir anschließend zwecks Sicherstellung der Güte der Quellen von unabhängigen Experten aus den entsprechenden Bereichen überprüfen lassen. Eine Übersicht der Quellen ist in Abschnitt B.1.1 zu finden. Bei der manuellen Suche konnten wir 104 Publikationen als relevant selektieren, die dann unser Startset für die SLR formten. Anschließend komplementierten wir unsere Ergebnisse mit Snowballing (vorwärts und rückwärts). „Beim Forward Snowballing werden neue Arbeiten auf der Grundlage der Arbeiten identifiziert, die die untersuchte Arbeit zitieren […]" [225]. Beim Backward Snowballing untersucht man die Referenzliste der zu untersuchenden Arbeit und prüft diese Referenz für Referenz auf Eignung, um diese in die Auswahl der Publikationen der SLR aufzunehmen. Sowohl bei der manuellen Suche, als auch beim Snowballing sind wir unserem Auswahlprozess (siehe Abschnitt B.1.3) gefolgt. Beide Suchen wurden von uns Dreien unabhängig voneinander durchgeführt und die Zuverlässigkeit der Auswahlverfahren wurde anhand der Fleiss' Kappa-Statistik [227] bewertet. Die Fleiss' Kappa-Statistik hilft bei der Berechnung des Grades der Übereinstimmung zwischen mehreren Prüfern. Der berechnete Wert von $\kappa = 0.81$ zeigte eine nahezu perfekte Übereinstimmung [228] bei der manuellen Suche, ebenso wie für das Snowballing mit einem κ von 0.87. Eine Gesamtübersicht über die Anzahl der in den verschiedenen Phasen der SLR geprüften und ausgewählten Publikationen findet sich in Abschnitt B.1.2.

Unser SLR-Prozess basierte teilweise auf dem von Wolfswinkel et al. [229] vorgeschlagenen Ansatz der Grounded Theory (GT) für Literaturrecherchen. Das Ziel bei diesem Ansatz besteht darin, eine detaillierte und sachbezogene Analyse eines Themenfeldes zu erreichen sowie dabei einige der Grundsätze der GT zu befolgen. Eine Literaturrecherche ist niemals vollständig, sondern höchstens gesättigt, so Wolfswinkel et al. [229]. Diese Sättigung ist erreicht, wenn sich aus den Daten, also den analysierten Publikationen, keine neuen Erkenntnisse oder Konzepte mehr ableiten lassen. Wir haben dieses Verfahren beim Snowballing angewandt und stellten bei der Begutachtung der letzten Publikationen in der ersten Iteration fest, dass wir die oben genannte Sättigung erreicht haben, denn gleich zu Beginn der zweiten Iteration konnten wir keine neuen Erkenntnisse und Konzepte mehr feststellen. Somit beendeten wir unsere SLR nach der ersten Iteration Snowballing, die in 125 Publikationen resultierte. Insgesamt lieferte unsere SLR somit 229 Publikationen.

Kodierung und Analyse Die Datenextraktion in einer SLR ist äußerst kritisch. Jede Publikation wurde gründlich analysiert, um alle relevanten Informationen in Bezug auf unsere Forschungsfragen zu extrahieren. Wir haben ein Formular zur Datenextraktion verwendet, um die extrahierten Informationen zu erfassen. Das Datenerfassungsformular erfasst die folgenden Informationen: (i) Metainformationen zur Publikation, wie Autoren, Titel, Veröffentlichungsjahr, Venue, Forschungsdisziplin, Suchmethode, (ii) Abdeckung der Forschungsfragen,

(iii) ob Ergbenisse durch Studien validiert wurden. Während der Datensynthese haben wir die extrahierten Informationen zusammengetragen und zusammengefasst, um Hinweise und Belege für die Beantwortung der vorgeschlagenen Forschungsfragen zu sammeln.

Für die qualitative Datenanalyse verfolgten wir einen Open-Coding-Ansatz [230]. Dieser Ansatz besteht aus bis zu drei aufeinanderfolgenden Kodierungszyklen. Für den ersten Zyklus haben wir *Initial Coding* [231] genutzt, damit die Ansichten und Perspektiven der Autoren im Code erhalten bleiben. Für den zweiten Kodierungszyklus setzten wir *Pattern Coding* [232] ein, d.h. wir gruppierten die Codes des ersten Zyklus nach Ähnlichkeiten und konnten diese somit kategorisieren. Diese Kategorien wurden ausführlich unter uns diskutiert, bis Konsens darüber herrschte, dass die Kategorien die eigentliche Bedeutung der Codes korrekt widerspiegeln. Diese entstandenen Kategorien halfen bei der Analyse, um Ordnung in die Daten zu bekommen und Muster in ihnen zu erkennen. Für den dritten Kodierungszyklus verwendeten wir *Protocol Coding* [233]. Der Kodierungsprozess deckt sich mit dem in Abschnitt 4.1.2.2 beschriebenen Prozess, so dass ich für Details an dieser Stelle dorthin verweise.

4.1.1.2 Datenvalidierung

Zur Validierung unserer in der Literatur gefundenen Daten, sowie dem Kodierungsprozess führten wie zwei Workshops durch. Ein Workshop ausschließlich mit Philosophen und Psychologen und einen ausschließlich mit Requirements Engineers. Der Workshop mit den Philosophen und Psychologen hatte die Zielsetzung, unsere Daten bezüglich unserer Definition für erklärbare Systeme (4.2) zu diskutieren und validieren. Der Workshop mit den Requirements Engineers zielte darauf ab, unsere Daten bezüglich Erklärbarkeit und den in Beziehung stehenden Qualitätsaspekten zu diskutieren und validieren. Weitere Informationen zu den Workshops und darin stattgefundenen Aktivitäten sind in Abschnitt B.1.4 zu finden.

4.1.1.3 Strukturierung des Wissens

Abschließend haben die unsere validierten Daten und das gewonnene Wissen aus den Schritten ① und ② zusammengetragen und strukturiert. Dazu haben wir zunächst das Konzept der Erklärbarkeit in einem Systemkontext operationalisiert, indem wir eine Definition des Konzepts herausgearbeitet haben. In dem Workshop mit Psychologen und Philosophen haben wir vorgeschlagene Definitionen mit denen aus der Literatur kombiniert. Die daraus resultierende Definition enthält mehrere Variablen, um so flexibel wie möglich an spezifische Projektkontexte angepasst werden zu können und gleichzeitig ein gemeinsames Verständnis von Erklärbarkeit unter den Beteiligten zu schaffen.

Darüber hinaus haben wie ein konzeptionelles Modell erstellt, um unseren Wissenskatalog zu gestalten. Dieses Modell veranschaulicht die Auswirkungen der Erklärbarkeit auf

verschiedene Qualitätsdimensionen [4, 165], die bei der Entwicklung von erklärbaren System berücksichtigt werden sollten. Der Wissenskatalog umfasst die 57 von uns identifizierten Qualitätsmerkmale, die durch Erklärbarkeit beeinflusst werden können, positiv oder auch negativ. Der vierte Beitrag dieser Forschungen ist ein Referenzmodell für Erklärbarkeit, das die vorgeschlagene Definition, das konzeptionelle Modell und den Wissenskatalog in einem Artefakt zusammenfasst. Dieses Referenzmodell soll Software Engineers bei der Entwicklung erklärungsfähiger Systeme helfen, indem es ihnen einen Leitfaden für die Hauptaspekte bietet, die während der drei wichtigsten Phasen des Software-Lebenszyklus zu berücksichtigen sind: Anforderungsanalyse, Entwurf und Bewertung.

Der Fokus dieser Dissertation liegt allerdings im Zusammenspiel von dem Konzept der Erklärbarkeit und Privatsphäre, daher seien die Forschungsbeiträge, des konzeptionelles Modells, des Wissenskatalogs und des Referenzmodells nur aus Gründen der Vollständigkeit erwähnt. Im weiteren Verlauf nutze ich daher lediglich die Definition für erklärbare Systeme (Abschnitt 4.2) als Artefakt für diese Dissertation.

4.1.2 Forschungsvorgehen - Modell der Beziehungen von Erklärbarkeit und Privatsphäre

Wie in Abschnitt 4.1 erwähnt, habe ich hier ebenfalls einen multimethodischen Ansatz gewählt. Auch dieser besteht aus einer systematischen Datenerhebung und einer qualitativen Analyse der Daten. Abbildung 4.2 zeigt einen Überblick über das von mir durchgeführte Forschungsvorgehen, das ich nachfolgend im Detail beschreiben werde.

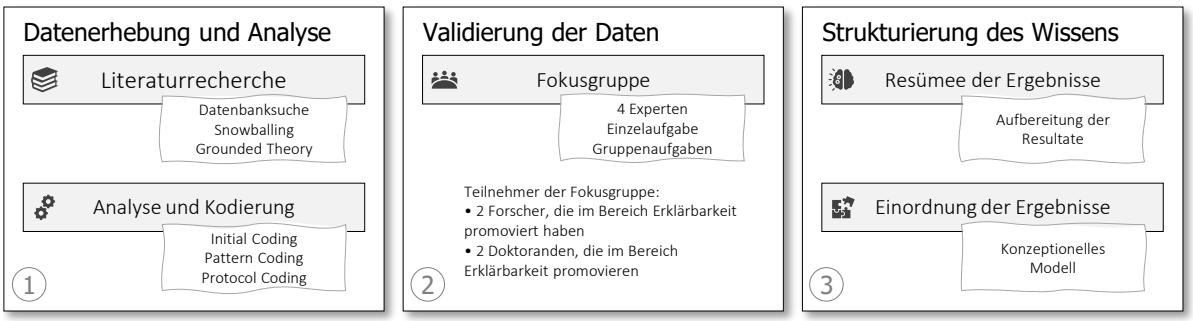


Abbildung 4.2: Übersicht des eigenen Forschungsvorgehens

4.1.2.1 Literaturrecherche

Um relevante Literatur zu identifizieren, die Aufschluss über das Zusammenspiel von Erklärbarkeit und Privatsphäre sowie deren Auswirkungen auf Softwarequalität gibt, habe ich mich für eine Datenbanksuche entschieden. Hierfür habe ich einen Such-String, basierend

auf für die Suche relevanten Schlüsselwörtern erzeugt. Die Schlüsselwörter hierfür, habe ich im Einklang mit anderen Forschern in gegenseitigen Diskussionen ausgewählt. Der folgende Such-String wurde für die Datenbanksuche verwendet:

> (*explain* ∨ *explanation* ∨ *explainability*) ∧
> ((*privacy* ∧ (*policy* ∨ *notice*)) ∨ „*privacy by design*“ ∨ „*privacy engineering*“) *-ai -xai*

Da besonders intensiv im Bereich der künstlichen Intelligenz und XAI an Erklärbarkeit geforscht wird, habe ich versucht Publikationen aus diesen Bereichen weitestgehend auszuschließen, da diese selten nur für meine Forschungen relevanten Bezug auf Softwarequalität und das Zusammenspiel von Erklärbarkeit und Privatsphäre haben. Daher ergeben sich die Parameter „*-ai -xai*“ im Such-String. Die Identifikation von Publikationen für eine LR ist ein zeit- und arbeitsintensiver Prozess. Daher erfolgte die Suche mit Google Scholar[2]. Google Scholar zeigt Ergebnisse aller wichtigen Datenbanken, wie z.B. ACM Digital Library, IEEExplore, Science Direct, Springer Link an. Bezugnehmend auf Yasnin et al. [234], die die Eignung von Google Scholar in Bezug auf Literaturrecherchen untersuchten, ist Google Scholar für eine Datenbanksuche geeignet, da 96% der Primärstudien im Vergleich zu herkömmlichen Datenbanksuchen gefunden werden konnten. Aufgrund zeitlicher Beschränkungen habe ich mich bei der Datenbank auf die ersten fünf Google Scholar-Ergebnisseiten beschränkt und die identifizierten Publikationen durch Snowballing komplementiert, um die Gefahr wichtige Literatur zu übersehen, abzumildern. Denn laut Wohlin [225] findet Snowballing sowohl Publikationen, die eine Datenbanksuche auch findet, so wie aber auch solche, die eine Datenbanksuche nicht findet.

Wie auch schon bei der in Abschnitt 4.1.1 beschriebenen SLR habe ich mich dem von Wolfswinkel et al. [229] vorgeschlagenem Ansatz der GT orientiert. Ich konnte bereits während der ersten Snowballing-Iteration eine Sättigung meiner Daten feststellen. Insgesamt führte meine LR zu 53 Publikationen.

Auswahlprozess der Publikationen Sowohl bei der Datenbanksuche als auch des Snowballing habe ich ein zweistufiges Auswahlverfahren angewandt, wie in Abbildung 4.3 dargestellt sowie meine zuvor definierten Einschluss- bzw. Ausschlusskriterien angewandt. In Phase 1 habe ich die in Frage kommenden Arbeiten anhand des Titels, der Zusammenfassung und der Keywords ausgewählt. In den Fällen, in denen diese Elemente keine ausreichenden Informationen lieferten, haben wir die Conclusion in unsere Überprüfung einbezogen, wie es Zhang und Babar [226] sowie Kitchenham [224] vorschlagen. Somit erhielt ich am Ende dieser

[2] `https://scholar.google.com`

ersten Phase eine Vorauswahl an möglichen Publikationen, die ich in einem nächsten Schritt genauer analysiert habe.

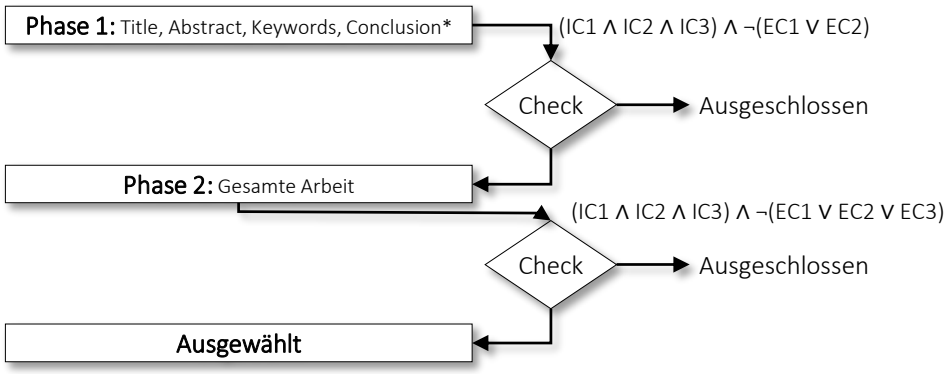


Abbildung 4.3: Übersicht des Auswahlprozesses meiner LR

Dazu habe ich in Phase 2 die vorausgewählten Arbeiten auf der Grundlage des Volltextes ausgewählt. Also den gesamten Inhalt des Papers studiert. Wenn mindestens ein Ausschlusskriterium erfüllt war, wurde die Studie nicht ausgewählt und aussortiert. Folgende Einschluss- und Ausschlusskriterien habe ich definiert:

IC_1 Adressiert das Thema Privatsphäre, sowie die Erklärung oder Übermittlung relevanter Informationen an den Endbenutzer

IC_2 Veröffentlicht zwischen 04.2002 – 12.2022

IC_3 *Peer-reviewed* Beitrag (Journal, Konferenz oder Workshop)

Ich habe Publikationen ausgeschlossen, die mindestens eines der folgenden Ausschlusskriterien erfüllten:

EC_1 nicht in englischer Sprache

EC_2 Tutorials, Proposals oder andere *non-peer-reviewed* Publikationen

EC_3 Arbeiten, die keinen Bezug zwischen Qualitätsaspekten sowie Privatsphäre und Erklärbarkeit herstellen

Ich habe April 2002 als Startdatum gewählt, weil hier die *Platform for Privacy Preferences (P3P) Specification* veröffentlicht wurde [235] (siehe Abschnitt 2.8.2). Durch die Wahl dieser Zeitspanne wollte ich einen möglichst breiten Überblick über das Thema gewinnen. Um eine Publikation aufzunehmen, müssen alle Einschlusskriterien erfüllt sein. Wenn mindestens eines der Ausschlusskriterien erfüllt war, wurde die Publikation abgelehnt. Formal ausgedrückt bedeutet das:

$$(IC_1 \wedge IC_2 \wedge IC_3) \wedge \neg(EC_1 \vee EC_2 \vee EC_3)$$

Da EC_3 nur auf den gesamten Inhalt einer Publikation anzuwenden ist, wurde es in der ersten Phase des Auswahlprozesses (Abbildung 4.3) nicht berücksichtigt.

4.1.2.2 Kodierung und Analyse

Um einen möglichst objektiven Kodierungsprozess zu gewährleisten und die Gefahr der eigenen subjektiven Färbung der Codes zu minimieren, habe ich für den Kodierungprozess Unterstützung von einer Forscherin (Larissa Chazette) erhalten, die im Bereich der Erklärbarkeit promoviert hat und über Hintergrundwissen im Bereich der Privatsphäre verfügt. Wie in Absatz 4.1.1.1 bereits beschrieben bestand unser Vorgehen ebenfalls von drei aufeinander folgenden Kodierungszyklen (siehe Abbildung 4.2). Dabei folgten wir wieder dem Open-Coding-Ansatz [230]. Im ersten Zyklus setzten wir *Initial Coding* [231] ein, um die Ansichten und Perspektiven der Autoren im Code zu erhalten. Beim zweiten Kodierungszyklus nutzten wir *Pattern Coding* [232]. Dadurch gruppierten wir die Codes des ersten Zyklus nach Ähnlichkeiten und konnten diese somit kategorisieren. Auch hier fanden Diskussionen bei Uneinigkeit statt, bis Konsens herrschte und beide der Ansicht waren, dass die Kategorien die eigentliche Bedeutung der Codes korrekt widerspiegeln.

Für die Analyse in Bezug auf die Beziehung zwischen Erklärbarkeit, Privatsphäre und anderen Qualitätsaspekten, verwendeten wir *Protocol Coding* [233] als dritten Kodierungszyklus. Hierfür diente uns die bereits existierende Liste von NFRs von Chung et al. [236]. Wurde eine Entsprechung zwischen einer NFR und einer Kategorie identifiziert, wurde der entsprechende Code zugewiesen. Wenn der Fall eintrat, dass keine solche Entsprechung mit der Liste von [236] identifiziert werden konnte, diskutierten wir und wiesen einen dem im Textfragment dargestellten Konzept entsprechenden Aspekt zu. Informationen zu den Codes sind in Abschnitt B.2 zu finden.

4.1.2.3 Validierung der Daten

Die Validierung der Ergebnisse aus Schritt ① (Abbildung 4.2) fand in einer Fokusgruppe stand. Diese bestand aus vier Teilnehmern. Zwei der Teilnehmer haben im Bereich der Erklärbarkeit promoviert, die zwei übrigen Teilnehmer sind Doktoranden und werden im Bereich der Erklärbarkeit promovieren. Zudem verfügen zwei der Teilnehmer über fundiertes Hintergrundwissen im Bereich der Privatsphäre.

Für das Treffen der Fokusgruppe war ein Zeitfenster von ca. 2,5 Stunden angesetzt und ich habe im Vorfeld eine vorbereitende Aufgabe an alle Teilnehmer gestellt, diese bis vor dem

Termin von allen Teilnehmern durchzuführen war. Zudem habe ich den Teilnehmern mit der Vorbereitungsaufgabe alle von mir identifizierten Qualitätsmerkmale mitgesendet, jedoch ohne Einordnung in die entsprechenden Dimensionen, um eine Beeinflussung der Teilnehmer zu verhindern. Bei der Vorbereitungsaufgabe sollten die Teilnehmer sich mit den Qualitätsdimensionen aus [4] vertraut machen. Anschließend sollten sich die Teilnehmer auf Basis Ihrer Expertise Gedanken machen, welche Qualitätsmerkmale Ihrer Meinung nach beim Zusammenspiel von Erklärbarkeit und Privatsphäre eine wichtige Rolle spielen und mit der mitgesendeten Liste abgleichen, ob eventuell Qualitäten fehlen oder Qualitäten dabei sind, welche es Expertensicht nicht dazugehören würden. Unterstützend dafür, habe ich den Teilnehmern zwei Szenarien zur Verfügung gestellt, die als Unterstützung für Ihre Überlegungen dienen sollten. Die Szenarien sind in Abschnitt B.2.1.2 zu finden.

Während des Treffens wurden zwei Aufgaben von der Gruppe bearbeitet. In der ersten Aufgabe sollten die Experten die Qualitätsmerkmale, ggf. ergänzt durch die von ihnen identifizierten, den einzelnen Dimensionen zuordnen. Jeder Teilnehmer wählte hierzu aus der alphabetisch sortierten Liste das nächste Qualitätsmerkmal aus und schlug eine Dimension vor, in die es eingeordnet werden sollte. Jeder Teilnehmer musste seine Entscheidung begründen und im Anschluss wurde diese mit der Gruppe diskutiert, bis Konsens erreicht wurde. Nachdem diese Aufgabe abgeschlossen war (Dauer: 1,5 Stunden), habe ich den Teilnehmern mein Modell mit den Qualitätsmerkmalen präsentiert. Anschließend habe wir die Unterschiede identifiziert und diskutiert, bis auch hier Konsens für jedes Qualitätsmerkmal erreicht wurde (Dauer: 1 Stunde). Auf Basis der Ergebnisse aus der Fokusgruppe, konnte ich mein nun validiertes Modell verfeinern und finalisieren.

4.1.2.4 Strukturierung des Wissens

Der letzte Schritt meiner Forschung bestand darin, das in den vorangegangenen Phasen gesammelte Wissen zu strukturieren und zu verwerten. Dazu habe ich das gewonnene Wissen aus der LR mit den Ergebnissen und Diskussionen der Fokusgruppe kombiniert. Das versetzte mich in die Lage, ein Modell der Beziehungen von Erklärbarkeit und Privatsphäre sowie deren Beziehungen zur Softwarequalität zu konstruieren. Das Modell hilft bei der Begutachtung der Auswirkungen vom Zusammenspiel zwischen Erklärbarkeit und Privatsphäre auf verschiedene Qualitätsdimensionen. Dieses geschaffene Artefakt, soll Software-Engineers bei der Entwicklung von privacy-aware Systemen unterstützen.

4.2 Definition für erklärbare Systeme

Eine rein abstrakte Definition von Erklärbarkeit ist für das RE nicht zweckdienlich, denn bevor Requirements Engineers den Bedarf für Erklärbarkeit eines Systems ermitteln können, müssen sie zuvor verstehen, was ein erklärbares System ausmacht. Daher wird eine Definition benötigt, die sich auf die Anforderungen an erklärbare Systeme konzentriert.

Erklärbarkeit ist mit der Offenlegung von Informationen verbunden, Das kann durch das Geben einer Erklärung erfolgen. Köhl et al. [52] vertreten die Auffassung, dass der Zugang zu Erklärungen ein System erklärbar macht. Was aber nun genau erklärt werden soll, bleibt hierbei allerdings offen. In der Literatur variieren die Definitionen von Erklärbarkeit in dieser Hinsicht erheblich. Darüber hinaus hat unsere Untersuchung weitere Aspekte aufgezeigt, in denen sich Definitionen von Erklärbarkeit unterscheiden. Folglich gibt es nicht *eine* Definition von Erklärbarkeit, sondern mehrere Komplementäre.

Auf Grundlage der Daten unserer SLR, komplementiert mit den Ergebnissen des Workshops mit Philosophen und Psychologen, haben wir eine Definition für erklärbare Systeme entwickelt, die je nach Projekt- oder Anwendungskontext angepasst werden kann.

Definition 4.2.1: Erklärbare Systeme

Ein System S ist in Bezug auf einen Aspekt X von S relativ zu einem Adressaten A im Kontext C genau dann erklärbar, wenn es eine Entität E (den Erklärer) gibt, die es A durch Angabe eines Informationskorpus I (die Erklärung von X) ermöglicht, X von S in C zu verstehen.

Unsere Definition fasst die für Requirements- und Software Engineers als relevant identifizierten Variablen eines erklärbaren Systems zusammen. Diese Variablen dienen als Orientierungshilfe für die Elemente, die in einem erklärbaren System von Bedeutung sind und daher während der Elizitierung und dem Entwurf berücksichtigt werden müssen.

In der Literatur gab es Unterschiede hinsichtlich der Werte der folgenden Variablen, die in der obigen Definition aufgeführt sind: *Aspekte eines Systems*, die erklärt werden sollten, *Kontexte*, in denen erklärt werden soll, die Entität, die erklärt (*der Erklärer*), und die *Adressaten*, die die Erklärung erhalten. Sich dieser Unterschiede bewusst zu sein, ist für Requirements Engineers von entscheidender Bedeutung, um die richtige Art von Erklärbarkeit für ein Projekt herauszufinden und die passenden Anforderungen an Erklärungen zu spezifizieren.

4.2.1 Aspekte eines Systems

Für die **Aspekte**, die erklärt werden sollen, haben wir in der Literatur die folgenden Facetten gefunden, die wir während unseres Workshops mit den Philosophen und Psychologen

validierten: das System im Allgemeinen (z.B. globale Aspekte eines Systems) [237] sowie etwas spezifischer dessen Argumentationsprozesse (z.B. Inferenzprozesse für bestimmte Probleme) [164]. Darüber hinaus die innere Logik des Systems (z.B., Beziehungen zwischen den Eingaben und Ausgaben) [52], die Interna des zugrunde liegenden Modells (z.B. Parameter und Datenstrukturen) [238], Intention (z.B. angestrebte Ergebnisse von Aktionen) [239] und auch das Verhalten (z.B, Aktionen in der realen Welt) [240]. Des Weiteren fanden wir die Entscheidung des Systems (z.B. zugrunde liegende Kriterien) [214], die Performanz (z.B. Vorhersagegenauigkeit) [241] und das Wissen des Systems über den Benutzer oder die Welt (z.B. Benutzerpräferenzen) [240].

4.2.2 Kontext und Erklärer

Ein **Kontext** wird durch eine Situation bestimmt. Diese Situation besteht aus einer Interaktion zwischen einem Benutzer, einem System, einer Aufgabe sowie einer Umgebung (englisch: *environment*) [242]. Hierbei können Zeitdruck, Anliegen an das System und die Art des Systems selbst mögliche Einflüsse auf den Kontext sein [243].

Erklärer beziehen sich auf ein System oder bestimmte Teile eines Systems, die schließlich die Stakeholder mit den benötigten Informationen versorgen. Semantisch gesehen erlaubt unsere Definition, dass diese spezifischen Teile des Systems nicht notwendigerweise technische Komponenten (wie Algorithmen oder sogar Hardware-Elemente) des Systems selbst sein müssen. In diesem Sinne könnte ein *explainer* auch eine Art Zwischeninstanz sein, eine Art externer *Mediator/Vermittler*. Dieser Vermittler fungiert als Schnittstelle zwischen dem System und dem Adressaten. Er erklärt „etwas“ und verhilft somit dem Adressaten, den entsprechenden Aspekt des Systems zu verstehen [243].

Die Definition ist bewusst offen mit Hinblick auf solche Fälle formuliert. Es obliegt der Person, die diese Definition anwendet zu entscheiden, wo die Grenzen eines erklärbaren Systems gezogen werden sollen. Im Rahmen dieser Dissertation liegt der Fokus auf selbsterklärenden Systemen. Also Systemen, die sich einem Endbenutzer direkt erklären. Um diesen Sachverhalt etwas greifbarer zu machen, folgt nun ein auf [3] basierendes Beispiel dazu.

> **Beispiel.** Ein Patient (Adressat) befindet sich in einem Krankenhaus und wurde von einem Arzt (Kontext) mit Hilfe eines medizinischen Diagnosesystems untersucht. Die medizinischen Befunde werden von dem System verarbeitet (Aspekte) und dem Patienten direkt in einem elektronischen Dashboard präsentiert. Diese Befunde können jedoch vom Patienten nicht direkt interpretiert und verstanden werden, da er nicht über das notwendige medizinische Fachwissen verfügt. Daher greift der Arzt als Vermittler ein und erklärt dem Patienten die Untersuchungsergebnisse in einer für den Patienten verständlichen Weise.

Dieses System könnte nach der vorgeschlagenen Definition als *erklärungsfähig* angesehen werden, da es die Ergebnisse an den Arzt übermittelt, der sie versteht, und der Arzt (Mediator) seinerseits in der Lage ist, dem Patienten die Ergebnisse des Systems auf der Grundlage der erhaltenen Erklärungen zu erläutern (erklären).

Wie oben beschrieben liegt der Fokus (Blickwinkel) meiner Dissertation auf selbsterklärenden Systemen. Das bedeutet für das gegebene Beispiel, dass das System unter Berücksichtigung dieses Blickwinkels nur dann als erklärbar gilt, wenn der Arzt der beabsichtigte Adressat der Erklärungen ist. Sind die Patienten hingegen die intendierten Adressaten, wäre das System (nur) dann erklärbar, wenn sich es dem Patienten direkt und umfassend erklärt, ohne dass ein Arzt als Vermittler eingreifen muss, um ein ausreichendes Verständnis auf Seiten des Patienten zu schaffen. Das könnte bedeuten, das gegebenenfalls keine medizinische Terminologie (Fachbegriffe) verwendet werden dürfen und die Untersuchungsergebnisse müssen für Laien klar und verständlich dargestellt werden. Daher ist die Zielgruppe entscheidet dafür, ob ein System als erklärbar angesehen werden kann oder auch nicht. Im Falle des obigen Beispiels bedeutet das, sollten die Ärzte als Zielgruppe, also die Endbenutzer, des medizinischen Diagnosesystems sein, gilt das System als erklärbar. Sind Patienten die Endbenutzer, wäre das System nicht als erklärbar anzusehen.

4.2.3 Verständnis des Adressaten

In der Literatur wird vielfach auf das hervorgerufene Verständnis der Adressaten als wichtigen Faktor für den Erfolg von Erklärbarkeit verwiesen, wie z.B. in [218, 237, 244, 245, 155]. Die Betrachtung von Erklärbarkeit im Kontext von Verständnis hat den Vorteil, dass Erklärbarkeit messbar wird. Es gibt eine Reihe an etablierten Methoden, um das Verständnis einer Person für etwas zu eruieren (z.B. A/B Tests [246, 247], Benutzerstudien [248, 249], Case Studies [250, 251], Fragebögen [252, 253], Interviews [241, 254] oder auch Usability-Tests [165]),

4.3 Erklärbarkeit und Privatsphäre

Die Qualität eines Systems und somit dessen Erfolg wird u.a. durch die bei der Entwicklung berücksichtigen NFRs sowie dessen Umsetzungen bestimmt [255]. „Obwohl jede NFR für sich genommen eindeutig und vom Stakeholder gewünscht sein kann, können [...] diese aufgrund inhaltlicher Zielkonflikte unter Umständen nicht zusammen [...]“ [73] realisiert werden. Dieses Verhalten kann als antagonistische Beziehung zwischen einzelnen Qualitäten bzw. Anforderungen angesehen werden [166, 256]. So kann von einem System beispielsweise verlangt werden, dass es sehr performant in Bezug auf Geschwindigkeit sein soll, gleichzeitig

soll es aber auch sehr sicher sein. In diesem Fall könnten aufwendige kryptografische Algorithmen die Systemperformanz beeinträchtigen. Somit hätten Performanz und Sicherheit hier eine antagonistische Beziehung.

Die beiden NFRs Erklärbarkeit und Privatsphäre (Privacy) pflegen ebenfalls eine antagonistische Beziehung [256]. Diese Erkenntnis konnten wir sowohl durch unsere SLR als auch durch den Workshop mit den Requirements Engineers bestätigen. So muss sorgfältig geprüft werden, *was* erklärt und *wie* erklärt werden soll. Denn Erklärungen können negative Einflüsse auf die Privatsphäre haben, da möglicherweise zu viele Informationen preisgegeben [245, 257, 258] oder auch für Erklärungen benötigt werden [259, 260]. Darüber hinaus besteht zudem ein potentiell negativer Einfluss zwischen Erklärbarkeit und der Privatsphäre von Unternehmen (in Bezug auf deren Geschäftsgeheimnisse) [161, 245, 166, 257]. Diese potentiell negativen Auswirkungen machen einmal mehr deutlich, wie wichtig es ist, Erklärbarkeit nicht als Allheilmittel, sondern als Mittel zum Zweck zu betrachten, um gewünschte andere Qualitätsziele zu erreichen und dieses beim Entwurf von Informationssystemen zu berücksichtigen.

Abbildung 4.4: Beziehungen von Erklärbarkeit und Privatsphäre mit anderen Qualitätsmerkmalen unter Einbeziehung der Dimensionen aus [4]

Den negativen Einflüssen stehen aber auch viele positive Einflüsse gegenüber, die mit einer Reihe anderer NFRs wiederum in Verbindung mit Privacy stehen und von essentieller Bedeutung sind. Abbildung 4.4 zeigt das konzeptuelle Modell der 32 Qualitätsaspekte, die bei

der Entwicklung von Systemen im Kontext von Erklärbarkeit und Privatsphäre berücksichtigt werden sollten. Das Modell basiert auf unserer gemeinsamen Arbeit [4] und ich habe es im Rahmen meiner eigenen Forschungen durch eine Literatursuche komplementiert, adaptiert und erweitert, um es gezielt auf den Kontext von Erklärbarkeit und Privatsphäre auszurichten. Das Modell enthält die essentiellen Qualitätsmerkmale, die von der Literatur genannt werden, wenn es um die Informierung von Endbenutzern im Umgang mit Ihrer Privatsphäre geht, also den Benutzern Hinweise oder Informationen irgendeiner Art (statische Text wie DSEs oder auch kurze Datenschutzhinweise) gegeben werden, um über Datenpraktiken eines Service-Anbieters in Kenntnis gesetzt zu werden. Berücksichtigt wurde hierfür Literatur, die im Kontext der Privatsphäre DSEs analysiert, Anforderungen an diese erhebt oder andere Verfahren vorschlägt, wie DSEs vereinfacht bzw. Nutzer in Bezug auf ihre Privatsphäre informiert werden können. Details zur Literatursuche und dem damit verbundenem Vorgehen sind in Abschnitt B.2 zu finden.

Die verschiedenen Qualitätsaspekte aus Abbildung 4.4 sind hierbei in die vier von uns identifizierten Qualitätsdimensionen (① Benutzerbedürfnisse, ② Kulturelle Werte, Gesetzten und Normen, ③ Domänenaspekte und Unternehmenswerte und ④ Projektbeschränkungen und Systemaspekte) [4] eingebettet. Das fünfte Element *Multidimensionale Qualitäten* bildet Fundament und Basis in Bezug auf Qualitätsziele in Verbindung mit Erklärbarkeit und Privatsphäre. Dieses Element kann als dimensionsübergreifend betrachtet werden, denn die darin enthalten Qualitäten haben wechselseitige Beziehungen mit mehr als einer der anderen Dimensionen. Understandability als Beispiel hat einen Einfluss auf die Dimension des Benutzers. Dieser soll in der Lage sein, eine erhaltene Erklärung zu verstehen. Dazu muss das System aber so entworfen worden sein, dass es als erklärbar angesehen werden kann, also in der Lage ist, sich dem Benutzer verständlich mitzuteilen.

4.3.1 Multidimensionale Qualitäten

Die multidimensionale Qualitäten sind, analog zu den *foundational* und *superordinated* Qualitäten, die wir in [4] identifiziert haben hier zusammen gefasst. Transparenz und Verständlichkeit bilden die Grundlage für alle vier Dimensionen und haben somit Einfluss auf die anderen Aspekte innerhalb dieser Dimensionen. Erklärungen eines Systems, dessen Prozesse und Ergebnisse können das Verständnis (**understandability**) auf vielen Ebenen erleichtern [261, 48, 262]. Außerdem tragen Erklärungen zu einer höheren Systemtransparenz (**transparency**) bei [263, 264, 265]. „Nutzer sollten in die Lage versetzt werden, die Risiken für die Privatsphäre zu verstehen“ [266, 267], was wiederum ihrerseits in Nachvollziehbarkeit (**comprehensebility**) mündet [268, 269, 186]. Das führt auf Seiten der Nutzer zu Vertrauen (**trust**) [41, 270, 271] und stärkt die Vertrauenswürdigkeit (**trustworthiness**)

des Anbieters [184, 272, 273]. In dieser Dissertation verwende ich das Qualitätsmerkmal (System-) **Transparenz** im Zusammenspiel von Erklärbarkeit und Privatsphäre in Bezug auf Privatsphäreaspekte, also das Offenlegen von Datenpraktiken, ihren Konsequenzen und Implikationen.

4.3.2 Kulturelle Werte, Gesetze und Normen

Gesetze und Normen gründen u.a. auf den kulturellen Werten einer Gesellschaft, weswegen diese symbiotische Beziehung sich in dieser Dimension widerspiegelt. Anforderungen an die Privatsphäre stützen sich auf gesetzliche Vorgaben und Regulierungen, sind also für Unternehmen rechtlich bindend (**accountability** [271, 273, 274], **lawful** [275, 276, 277] und **compliance** [276, 278, 279]) und folgen **ethischen** [263, 280, 281] und **fairen** [275, 277, 282] Prinzipien im Umgang mit den Daten [283, 122].

4.3.3 Benutzerbedürfnisse

Aus der Sichtweise des RE zählen (End-) Benutzer ebenfalls zu einer wichtigen Klasse von Stakeholdern [67]. In der Regel sind die Benutzer nicht mit den technischen Details und der Funktionsweise der von ihnen verwendeten Systeme vertraut [243] und verfügen meist auch nicht über fundiertes IT Wissen [284].

Wenn Erklärbarkeit in ein System „integriert“ wird, haben verschiedene Gruppen von Benutzern auch unterschiedliche Erwartungen, Erfahrungen, persönliche Werte, Vorlieben und Bedürfnisse. Das bedeutet, dass Individuen Qualität unterschiedlich wahrnehmen können. Gleichzeitig beeinflusst die Erklärbarkeit Aspekte, die aus der Sicht der Nutzer äußerst wichtig sind. Das ist ebenfalls im Kontext der Privatsphäre der Fall. Es existiert eine ganze Reihe an Forschung zu den unterschiedlichen Einstellungen und Haltungen (privacy attitudes) von Endbenutzer in Bezug auf Ihre Privatsphäre [192, 111, 285, 112, 109, 286].

Es ist wichtig, dies bereits beim Entwurf und Design von Software-Systemen zu berücksichtigen [49, 140]. Auf oberster Ebene, kann die **User Experience** hier stark profitieren [282, 109, 114]. Verständliche und nachvollziehbare Informationen zum Datenschutz wirken sich positiv auf die **Privacy Awareness** [111, 287, 189], den empfundenen Nutzen (**perceived value**) [49, 164, 109] und die empfundene Nützlichkeit (**perceived usefulness**) [192, 189, 288] aus. Dadurch ermöglicht man dem Benutzer die Kontrolle über seine eigenen Daten (**user control**) [270, 289, 290] zu behalten und befähigt ihn informierte Entscheidungen (**decision making**) [164, 266, 48] in Bezug auf die Privatsphäre zu treffen. Das führt wiederum zu deutlich höherer Zufriedenheit (**user satisfaction**) [49, 140, 287] sowie Akzeptanz auf Seiten der Nutzer (**user acceptance**) [281, 288, 291] und schließt somit den Kreis zur User Experience.

4.3.4 Domänenaspekte und Unternehmenswerte

Diese Dimension wird von zwei Aspekten geprägt: (i) den Unternehmenswerten sowie der Vision einer Organisation [292] und (ii) den Domänenaspekten, die das Design eines Systems prägen, da Erklärungen in einigen Domänen dringlicher sein können als in anderen. Da i.d.R. jedes Unternehmen Daten ihrer Kunden verarbeitet oder auch sammelt, ist es von entscheidender Bedeutung wie Unternehmen ihre Informationen zur Sicherheit und Datenpraktiken darstellen. Denn es beeinflusst das Vertrauensverhältnis, welches Nutzer gegenüber diesem Unternehmen aufbauen (**customer trust**) [184, 281, 288]. Dieses gewonnene Vertrauen wirkt sich wiederum positiv auf die Kundenbindung/-treue (**customer loyalty**) [271, 184, 281] aus. Vertrauen ist hierbei als ein Mediator zwischen „der wahrgenommenen Sicherheit im Umgang mit privaten Daten und der Loyalität gegenüber“ eines Service Providers anzusehen [271]. Die Verlässlichkeit (**reliability**) [273, 278, 293], ob Unternehmen sich an Gesetze und Normen im Umgang mit persönlichen Daten halten und ob diese Informationen transparent kommuniziert werden, ist ein weiterer wichtiger Faktor in dieser Dimension. Denn rechtskonformes Design der Systeme „kann eine zuverlässige und transparente Infrastruktur für die Einbettung relevanter rechtlicher Schutzmaßnahmen in die Benutzeroberflächen, Datenschutzrichtlinien und Nutzungsbedingungen von Produkten und Dienstleistungen […] bieten“ [273]. Die Verlässlichkeit ist hierbei stark mit der Vertrauenswürdigkeit (trustworthiness) verwoben. Transparenz und Offenheit im Umgang mit Daten sollten also keine Bedrohung für ein Unternehmen darstellen, sondern vielmehr ein Chance, Vertrauen zu schaffen und Kunden zu binden.

4.3.5 Projektbeschränkungen und Systemaspekte

Personen, die Systeme designen, erstellen und programmieren, sind unter anderem Entwickler, Qualitätsingenieure und Softwarearchitekten. Sie zählen ebenfalls zu den Stakeholdern [294], da ohne sie die Systeme gar nicht erst existieren würden. Im Allgemeinen verfügen Vertreter dieser Gruppe über ein hohes Fachwissen über die Systeme und haben ein starkes Interesse an deren Entwicklung und Verbesserung. Diese Dimension wird durch zwei Aspekte geprägt: Projektbeschränkungen und Systemaspekte. Bei den Projektbeschränkungen handelt es sich um die nicht-technischen Aspekte eines Systems [295], während sich die Systemaspekte eher auf interne Aspekte des Systems beziehen, wie z. B. Leistung und Wartbarkeit. Im Kontext von Erklärbarkeit und Privatsphäre sind darüber hinaus auch weitere Systemaspekte von Relevanz. Zugänglichkeit (**accessibility**) [48, 282, 118] ist beispielsweise besonders in der heutigen Zeit sehr wichtig, da wir in vielen Situationen des Alltags von Software umgeben sind und damit keine Personen oder Minderheiten ausgeschlossen werden [296]. Im Kontext der Privacy

bedeutet dies zudem, dass Informationen zu Datenpraktiken sichtbar und leicht zugänglich gemacht werden müssen. Zudem sollten hier adaptive Mechanismen entwickelt werden (**adaptability**) [266, 272, 187]. Denn „kontext-adaptive Datenschutzmechanismen können den Nutzern helfen, die Auswirkungen ihrer Handlungen und Entscheidungen auf die Privatsphäre zu verstehen, wenn sie mit […] Systemen interagieren“ [187] und das führt zu Wahlmöglichkeiten durch die diese Informationen **actionable** werden [41, 184, 272]. Vor diesem Hintergrund zeigt sich dann auch die Bedeutung des Kontextbezugs (**contextual**) [266, 192, 291]. Die Entscheidung, die persönlichen Daten zu teilen, kann nämlich stark vom Kontext abhängen. Betrachten wir dazu folgendes Beispiel in Anlehnung an [266]:

> Alice wohnt in einem Smart Home, leidet unter einer chronischen Nierenerkrankung und hat zu Hause ein Hämodialysegerät zur Therapie. Nun möchte der örtliche Stromversorger gerne Verbrauchsdaten der Bewohner sammeln, um möglichst ökologisch und ökonomisch den Strom in der Stadt zu verteilen. Darüber hinaus kann er den Verbrauchern so auch detaillierte Energiespartips geben, damit diese Kosten und Verbrauch reduzieren können. Grundsätzlich ist Alice an den Verbrauchertips des Stromversorgers interessiert und ist bereit, feingranulare Daten über ihren Stromverbrauch mit dem Versorger zu teilen. Allerdings möchte Alice das nicht, wenn sie ihr Hämodialysegerät verwendet, da sie ihrem Stromversorger keine Auskünfte über ihren Gesundheitszustand zukommen lassen möchte.

Damit Alice nun aber derartige Entscheidungen treffen kann, müssen die Systeme ausreichend Informationen bzw. Mechanismen bereitstellen (**informativeness**) [262, 283, 175], die vollständig, korrekt (**correctness**) [275, 291, 297] und präzise (**precision**) [262, 186, 140] (implementiert) sind. So entstehen schließlich effiziente (**efficiency**) [186, 278, 282] und praktikable Lösungen.

4.4 Einschränkungen und Bedrohung der Validität

Die nun folgenden Erläuterungen zu den Einschränkungen und zur Bedrohung der Validität beziehen sich auf beide hier in diesem Kapitel angewandten Forschungskonzepte. Beide Konzepte verwenden ähnliche Ansätze und werden daher gemeinsam behandelt. Daher verwende ich das Wort „wir“, wenn es beide Forschungsvorgehen betrifft, ansonsten das Wort „ich“. Die hier angewandten Forschungsvorgehen (Abschnitt 4.1.1 und auch Abschnitt 4.1.2) basieren ausschließlich auf einer qualitativen Datenanalyse. Dadurch ist die Möglichkeit gegeben, dass abgeleitete Ergebnisse durch Subjektivität der Forscher im Zuge des Analyse beeinflusst wurden. Um dieser Gefahr zu begegnen, haben wir uns für einen multimethodischen Ansatz entschieden. Somit konnten Ergebnisse erzielt werden, die robuster und belastbarer sind, als

bei Studien mit nur einer Methode. Nachfolgend erörtere ich die wichtigsten Bedrohungen, die die Validität unserer Forschung bedrohen können.

(S)LR und Kodierung Der Recherche-Prozess bei einer SLR bzw. LR setzt ein gemeinsames Verständnis aller Forscher, hinsichtlich der verwendeten Such- und Analysemethoden voraus. Ein Missverständnis der Methoden und Konzepte birgt die Gefahr, dass Ergebnisse verfälscht werden könnten. Abgemildert haben wir diese Gefahr, indem wir ein Überprüfungsprotokoll erstellt und es zur Beginn der Literatursuche besprochen haben, um ein gutes gemeinsames Verständnis zu erreichen. Zudem haben wir Ein- und Ausschlusskriterien formuliert, um Verzerrungen aufgrund subjektiver Entscheidungen im Auswahlprozess zu verringern. Dabei sind einige dieser Kriterien objektiv, wie z.B. der Zeitraum, andere aber, z.B. den Inhalt betreffend immer noch subjektiv. Um das Problem der Voreingenommenheit zu verringern, haben wir die Analyse unabhängig durchgeführt. Gleiches gilt hier für die durchgeführten Kodierungsprozesse. D.h. sowohl bei der Literaturrecherche als auch bei der Kodierung wurde die Entscheidung über die Aufnahme oder den Ausschluss (für ein Paper) bzw. die Code-Zuweisung (für die extrahierten Daten) bei Uneinigkeit von allen Forschern getroffen und durch die Fleiss' Kappa-Statistik validiert.

Modell der Beziehungen von Erklärbarkeit und Privatsphäre Die Gruppierung und Kategorisierung der Qualitätsaspekte in ihre verschiedenen Dimensionen war anfällig für eigene subjektive Verzerrungen. Um dies abzumildern, habe ich mich bei der Kategorisierung auf bekannte Konzepte aus der Literatur gestützt und die Fokusgruppe mit Experten aus diesem Bereich durchgeführt. Somit konnte ich die Gruppierung und Kategorisierung durch interne und externe Validierung kontrollieren. Bei den internen Überprüfungen wurde die Kategorisierung unter den zwei Forschern diskutiert. Dadurch konnten Unklarheiten geklärt und Einigung erzielt werden. Bei den externen Überprüfungen habe ich die Erkenntnisse aus der Literatur mit dem Expertenwissen verglichen. Aufgrund dieses Vorgehens bin ich zuversichtlich, ein angemessenes Maß an Validität des Modells erreicht zu haben.

4.5 Fazit – Erklärbarkeit, Privatsphäre und Softwarequalität

Intuitiv hat die Erklärbarkeit einen großen Einfluss auf die Privatsphäre. Wenn Alice persönliche Informationen (Gefühle, Gedanken oder auch Daten wie Geburtsdatum etc.) über sich preisgegeben soll, hängt die Herausgabe dieser mit Sicherheit von verschiedenen Variablen bzw. Fragen ab: (i) wer möchte das wissen, (ii) warum ist das wichtig, (iii) vertraue ich der Situation, usw. Fragt zum Beispiel eine vertraute Person wie Bob, Alice nach diesen Informationen, stellt die Herausgabe der Informationen womöglich kein größeres Problem für Alice

dar. Fragt jemand Fremdes, könnte es schon anders aussehen und Alice würde gerne wissen, warum die Herausgabe wichtig sei. Alice möchte also vor der Herausgabe der Information eine Erklärung haben. Wie an diesem Beispiel zu sehen, sind hier eine ganze Reihe von Aspekten wie Vertrauen, Vertrauenswürdigkeit, Verständnis, Nachvollziehbarkeit und Transparenz involviert, die ebenfalls notwendig sind zu berücksichtigen, wenn wir mit Informationssystemen interagieren. Während der Interaktion mit diesen Systemen sind zudem noch eine Reihe weiterer Qualitätsaspekte involviert, die bereits beim Entwurf von Informationssystemen durch Software Engineers berücksichtigt werden sollten, wie das Modell in Abbildung 4.4 veranschaulicht, welches einen wichtigen Teil zur Beantwortung von RQ1 beiträgt.

> **Beantwortung RQ1:** Die Beziehung von Erklärbarkeit und Privatsphäre kann von antagonistischer Natur sein, wenn Erklärungen z.B. ungewollt private Informationen preisgeben. Im beiderseitigem Zusammenspiel beeinflussen Erklärbarkeit und Privatsphäre eine Reihe unterschiedlicher Qualitätsaspekte in verschiedenen Dimensionen, was in Abbildung 4.4 dargestellt ist, und sie legen das Fundament für Transparenz, Verständnis, Aufklärung, Vertrauen und Vertrauenswürdigkeit in und von Informationssystemen.

Die Antwort von RQ1 unterstreicht die Wichtigkeit des Konzepts der Erklärbarkeit im Kontext der Privatsphäre. Aber wie kann das gewonnene Wissen um das Konzept der Erklärbarkeit nun für die Entwicklung datenschutzfreundlicher Systeme eingespannt werden? Dieser Frage gehe ich in den folgenden Kapiteln auf den Grund.

5

Unterstützung der Endbenutzer beim Verständnis von Datenschutzerklärungen

Bei jeder Nutzung eines digitalen Dienstes oder beim „Betreten“ einer Internetseite gehen wir automatisch und implizit eine Art Vertrag zwischen dem Anbieter des Dienstes bzw. der Internetseite sowie der *Freigabe* unserer persönlichen Daten ein. Meist geschieht dies unwissentlich. Hierbei stellt die reine Nutzung der privaten Daten *a priori* keine Bedrohung der eigenen Privatsphäre dar. Allerdings hängt dieses primär vom Unternehmen selber ab, welche Daten seiner Anwender es sammelt, verarbeitet und speichert. Die rechtlichen Rahmenbedingungen hierzu werden durch Gesetze und Normen abgesteckt, die von regulatorischen Instanzen (Regierungen oder Staatenverbunde wie die EU) festgesetzt und vorgeschrieben werden, somit also rechtlich bindend für Unternehmen werden. Der primäre und meist einzige Kanal, durch den Anwender das „Kleingedruckte“ dieses oben angesprochenen Datenüberlassungskontraktes finden, sind die von den Unternehmen verfassten Datenschutzerklärungen (DSEs). Diese orientieren sich an den gesetzlichen Vorgaben und haben das Ziel, Nutzer über Datenpraktiken des jeweiligen Unternehmens aufzuklären.

Allerdings verfehlen DSEs in der Praxis dieses Ziel. DSEs sind sehr wortreich und langatmig [45, 46], schwer zu verstehen [47, 48], da Sie meist in einer für den Endbenutzer

schwer verständlichen (Fach-) Sprache verfasst wurden [175, 298], es zu lange dauert, die Texte zu lesen [48, 109] und sie auch nicht immer leicht zugänglich sind [45, 270]. Als Folge gelten DSEs als das am wenigsten gelesene Dokument einer Internetseite [48] und werden von den Benutzern meist ignoriert [43, 44], ungeachtet ihrer Bedenken bezüglich des Datenschutzes [51, 48].

Daher sind neue Ansätze erforderlich, um den Endbenutzern verständliche Informationen zum Schutz der Privatsphäre zur Verfügung zu stellen. Es existieren unterschiedliche Ansätze und Bemühungen der Forschung, DSEs zugänglicher und für den Benutzer verständlicher zu machen [49, 45, 262, 265, 299]. Allerdings sind viele davon theoretische Ansätze oder erfordern die Nutzung eines bestimmten Service Dritter. Ein Überblick zu verwandten Arbeiten dieses Themenbereichs ist in Abschnitt 2.8.2 zu finden.

Der hier vorgestellte Ansatz zur besseren Verständlichkeit von DSEs, hatte u.a. die Zielsetzung, praktisch und einfach nutzbar sowie direkt in der gewohnten Arbeitsumgebung eines Endbenutzers eingebettet zu sein. Wir verknüpfen unseren Ansatz mit dem Konzept der Erklärbarkeit. Durch die Integration von Erklärungen in Systeme ist es zum Beispiel möglich, Informationen unter anderem über das Systemverhalten selbst, die Entscheidungsfindung oder auch interne Parameter des Systems anzubieten, also eine Form der Transparenz zu schaffen. Aus diesem Grund könnten Erklärungen auch ein effizientes Mittel sein, um die Nutzer über Datenschutzaspekte zu informieren und sie in die Lage zu versetzen, bewusstere Entscheidungen in Bezug auf das System zu treffen, dem sie vertrauen wollen. Nachfolgend bezeichne ich das entwickelte Analysewerkzeug für DSEs als *PriX - **Pri**vacy Policy e**X**plainer.*

Zugehörige Publikationen Der in diesem Kapitel vorgestellte Forschungsbeitrag entstand in Kollaboration[1] mit fünf anderen Forschern: Larissa Chazette, Jil Klünder, Lukas Köhler, Kai Korte und Kurt Schneider. Dieses Kapitel gründet auf der gemeinsamen Zusammenarbeit und die Ergebnisse dazu wurden in [8], [42] und [118] veröffentlicht, worauf dieses Kapitel basiert.

5.1 Forschungsvorgehen

Unser Forschungsvorgehen besteht aus drei Hauptphasen: ① Konzept und Planung, ② Implementation und ③ Evaluation durch eine Benutzerstudie (siehe Abbildung 5.1). Die einzelnen Phasen werden im Folgenden genauer beschrieben.

[1] Aufgrund der Kollaboration mit meinen Forschungskollegen verwende ich das „wir“.

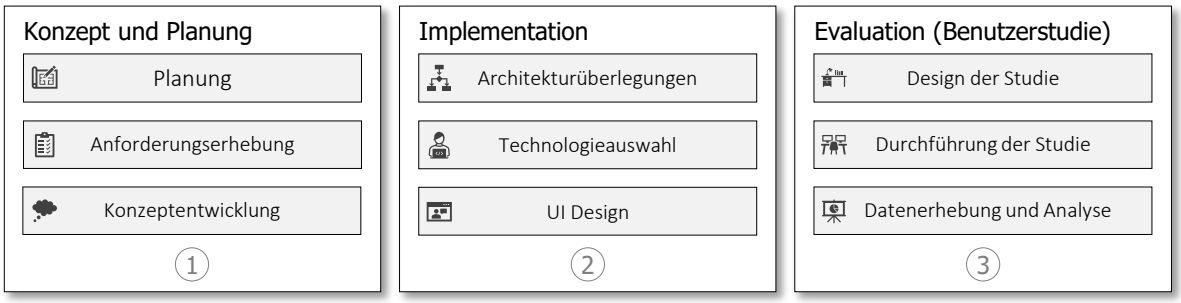


Abbildung 5.1: Übersicht des Forschungsvorgehens

5.1.1 Konzept und Planung

Zu Beginn haben wir verwandte Literatur (Abschnitt 2.8.2) herangezogen und einige der existierenden Ansätze zur Aufbereitung von Datenschutzerklärungen [49, 262, 265, 300] als Grundlage für unser weiteres Vorgehen genutzt. Da das Analysewerkzeug möglichst nahtlos in die Arbeitsumgebung eines Endbenutzers eingebettet werden sollte, entschieden wir uns für eine Erweiterung, die in den jeweiligen Browser integriert werden kann (Browser-Erweiterung) und eine spätere Evaluation durch eine Benutzerstudie.

5.1.1.1 Anforderungserhebung

Der Fokus von PriX liegt auf der visuellen Aufbereitung, die dem Endbenutzer einen schnellen Überblick über die Datenpraktiken eines Service-Anbieters liefert und gleichzeitig Endbenutzer in Bezug auf ihre privaten Daten sensibilisieren soll. Darüber hinaus soll die Benutzung der Software möglichst geringe Computer-Kenntnisse und kein fundiertes Wissen über den Datenschutz erfordern. Zu diesem Zweck wurden folgende, in Tabelle 5.1 zu findende, grundlegende Anforderungen ermittelt.

5.1.1.2 Konzeptentwicklung

Das zugrunde liegende Konzept hinter PriX lässt sich mit dem in Abbildung 5.2 dargestellten Schaubild beschreiben.

Abbildung 5.2: Übersicht des Konzepts

Als erstes wird die Internetseite, die der Benutzer aktuell geöffnet hat automatisiert auf das Vorhandensein einer DSE (Absatz 5.1.1.2) geprüft. Hat PriX eine DSE gefunden, wird diese

Tabelle 5.1: Übersicht der grundlegenden Anforderungen an das Analysewerkzeug für DSEs nach Köhler [8]

Anforderung	Beschreibung
R1	Die Anwendung soll einfach erlernbar sein.
R2	Die Anwendung muss einfache Möglichkeiten bieten, einen Überblick über eine DSE zu erhalten.
R3	Die Anwendung muss, soweit möglich, autonom arbeiten, um die Anzahl von Benutzerinteraktionen zu minimieren.
R4	Die wichtigen Informationen der DSE sollten klar erkennbar und schnell erreichbar sein.
R5	Da die Anwendung als Browser-Erweiterung entwickelt werden soll, ist ein modularer Aufbau der einzelnen Funktionen von besonderer Wichtigkeit.
R6	Die Analyse einer DSE soll die Inhalte korrekt wiedergeben.
R7	Die Anwendung soll portabel und in verschiedenen Browsern lauffähig sein.

analysiert (Absatz 5.1.1.2) und anschließend visuell für den Benutzer zwecks vereinfachtem Verständnis aufbereitet (Absatz 5.1.1.2).

Auffinden von Datenschutzerklärungen Es existiert kein einheitlicher Standard oder einheitliches Format für die Struktur oder den Aufbau einer DSE. Darüber hinaus ist auch nicht spezifiziert, wie eine Internetseite, die eine DSE anbietet, diese einbetten bzw. verlinken muss. Um diese Herausforderung zu meistern, haben wir 100 der am häufigsten besuchten Internetseiten im Internet[2] manuell überprüft. Wir wählten diese Liste aus, um sicherzustellen, dass wir Internetseiten untersuchen, die im täglichen Surfverhalten der Endnutzer eine hohe Relevanz haben. Ziel unserer Untersuchung war es, Gemeinsamkeiten in der Art und Weise zu ermitteln, wie eine DSE strukturiert ist und wie Internetseitenbetreiber diese in ihre Internetseiten einbetten. Das Ergebnis hinsichtlich der Verlinkung von DSEs war, dass bei allen untersuchten Internetseiten das Wort „*privacy*“ im Verzeichnispfad der URL[3] oder in der Subdomain enthalten war. In seltenen Fällen kann es vorkommen, dass Internetseiten mehr als eine DSE haben, z.B. wenn ein Unternehmen mehr als ein Produkt verkauft und diese auf unterschiedliche DSEs verweisen. In diesen Fällen hat der kürzeste Link (Anzahl der Zeichen in der URL) in der Regel zu einer allgemeinen Datenschutzerklärung geführt. Dies wurde bei der Entwicklung des Algorithmus, der automatisch nach der Datenschutzerklärung sucht, berücksichtigt.

Was den Inhalt einer Internetseite betrifft, so hat unsere Untersuchung ergeben, dass DSEs in der Regel eine bestimmte Gruppe von Wörtern enthalten, die wir weiterhin als *Policy-Indikatoren* bezeichnen. Typische Policy-Indikatoren sind Wörter bzw. Begriffe wie *„privacy notice“*, *„privacy statement“*, *„privacy policy“*, *„GDPR“*, *„your privacy“* oder *„data protection“*.

[2] `https://moz.com/top500`, Stand vom: 15.07.2021

[3] URL steht für Uniform Resource Locator [301]

Auf der Grundlage dieser Erkenntnisse sind wir in der Lage, die DSE einer Internetseite mit einem sehr hohen Grad an Genauigkeit automatisiert zu finden. Formal ausgedrückt, ist dieses Vorgehen in Algorithmus 1 und Algorithmus 2 zu finden.

Algorithmus 1 Suche Hyperlinkkandidaten zu einer Datenschutzerklärung

```
1  function getPrivacyURL
2      links ← fetchAllURLs()
3      privacyLinks ← new Dictionary()
4      for link ∈ links do
5          if link.text.contains("privacy") then
6              privacyLinks.add(link.hash(), link.href)
7          end if
8      end for
9      privacyPolicyLink ← getShortest(privacyLinks)
10     return privacyPolicyLink
11 end function
```

Zeile 2 der Funktion `getPrivacyURL` aus Algorithmus 1 sammelt alle Hyperlinks der aktuell besuchten Internetseite. In der for-Schleife in Zeile 4 wird geprüft ob das Wort „privacy“ im Linkpfad enthalten ist und ggf. zum Dictionary hinzugefügt. In Zeile 9 wird dann über die Funktion `getShortest` der kürzeste Linkpfad des Dictionarys gesucht und in Zeile 10 als Rückgabewert zurückgegeben.

Algorithmus 2 URL-Prüfung auf Datenschutzerklärung

```
1  function isPrivacyPolicy(url)
2      indicators ← 0
3      for policyIndicator ∈ policyIndicators do
4          if getTitleTag(url).text.contains(policyIndicator) then
5              indicators ← indicators +1
6          end if
7          for tag ∈ getAllHeadingTags(url) do
8              if tag.text.contains(policyIndicator) then
9                  indicators ← indicators +1
10             end if
11         end for
12     end for
13     return indicators ≥ needed                          ▷ needed ← 2
14 end function
```

Ist nun der Hyperlinkkandidat, der zur möglichen DSE der Internetseite führt identifiziert, wird dieser nun auf das Vorhandensein der oben beschriebenen Policy-Indikatoren überprüft (siehe Algorithmus 2). Die Variable `indicators` in Zeile 2 zählt die Anzahl der gefundenen Policy-Indikatoren. Diese muss größer gleich zwei sein, da die Indikatoren sowohl im

Titel (Zeile 4) als auch in den Überschriften-Tags (<h1> bis <h6>, Zeile 8) der Internetseite vorkommen müssen.

Analyse von Datenschutzerklärungen PriX nutzt ML für die Analyse und der *OPP-115 Korpus* [302] dient hierfür als Trainings-Set. OPP-115 steht hierbei für Online Privacy Policies, set of 115. Das ist eine Sammlung von real existierenden DSEs, die in natürlicher Sprache vorliegen (amerikanisches Englisch). Hierbei wurde jede dieser DSEs von drei graduierten Studenten der Rechtswissenschaften gelesen und annotiert. Die einzelnen Abschnitte der DSEs wurden von den Studierenden in die zehn Kategorien unterteilt, die in Tabelle 5.2 zu finden sind. Der Datensatz wurde vom Usable Privacy Project[4] für Lehr- und Forschungszwecke zur Verfügung gestellt.

Tabelle 5.2: 10 Datenkategorien einer DSE, in Anlehnung an [8]

Kategorie	Beschreibung
First Party Collection	Informationen über Daten, die vom Betreiber der Internetseite gesammelt und wie diese genutzt werden.
Third Party Sharing/Collection	Informationen darüber, ob Daten mit Dritten getauscht werden und in welchem Umfang dies geschieht.
Data Security	Informationen darüber, ob und wie vom Betreiber gespeicherte Daten geschützt werden.
User Access, Edit and Deletion	Informationen darüber, wie der Nutzer auf seine Daten zugreifen und diese bearbeiten oder löschen kann.
User Choice/Control	Informationen darüber, ob und wie weit der Nutzer selber entscheiden kann, welche Daten über ihn gespeichert werden dürfen.
Data Retention	Informationen darüber, wie und wie lange Daten seitens des Betreibers der Internetseite gespeichert werden.
Policy Change	Informationen darüber, wie Veränderungen in der Datenschutzerklärung veröffentlicht werden.
Do Not Track (DNT)	Informationen darüber, wie mit DNT-Anfragen umgegangen wird.
International and Specific Audiences	Informationen für Besucher aus dem Ausland und spezifische Personengruppen.
Other	Andere Informationen, wie beispielsweise Möglichkeiten zur Kontaktaufnahme mit dem Betreiber der Internetseite.

Für jede dieser zehn Kategorien haben wir separat einen Klassifizierungsalgorithmus trainiert. Damit stellen wir eine möglichst genaue Einordnung der Sätze einer DSE in die entsprechende Kategorie sicher. Als Klassifizierungsalgorithmen kommen Naive Bayes [303] und Random Forest [304] zum Einsatz. Zu Beginn des Trainingsprozesses und zur Vorbereitung der Daten für die Klassifizierungsalgorithmen werden die Daten zunächst mithilfe

[4] https://usableprivacy.org

eines Stemming-Algorithmus [305] bearbeitet und nachfolgend findet eine Feature-Extraction mithilfe eines CountVectorizers [306] statt. Anschließend wird durch Verwendung des Tf-idf-Maßes [307] den extrahierten Features eine bestimmte Relevanz zugeordnet. Die Daten werden abschließend den Klassifizierern übergeben, damit diese die jeweiligen Trainingsmodelle erzeugen können. Je nach Performanz des Klassifizierers, wird der besser abschneidende Algorithmus für die jeweilige Datenkategorie schließlich bei der Analyse von DSEs eingesetzt.

Visuelle Darstellung Die analysierten Informationen müssen nun dem Nutzer in angemessener Weise präsentiert werden. Da der Nutzer die DSE verstehen soll, soll PriX eine Erklärung liefern, um den Nutzer über Datenpraktiken zu unterrichten. Nach der Definition von erklärbaren Systemen (Abschnitt 4.2) ist ein System erklärbar, wenn es einem Adressaten A eine Erklärung I über einen bestimmten Systemaspekt X präsentiert, so dass A diesen X des Systems verstehen kann. Ausgehend von dieser Definition definieren wir eine *Visual Explanation* (*VE*) wie folgt:

Definition 5.1.1: Visual Explanation

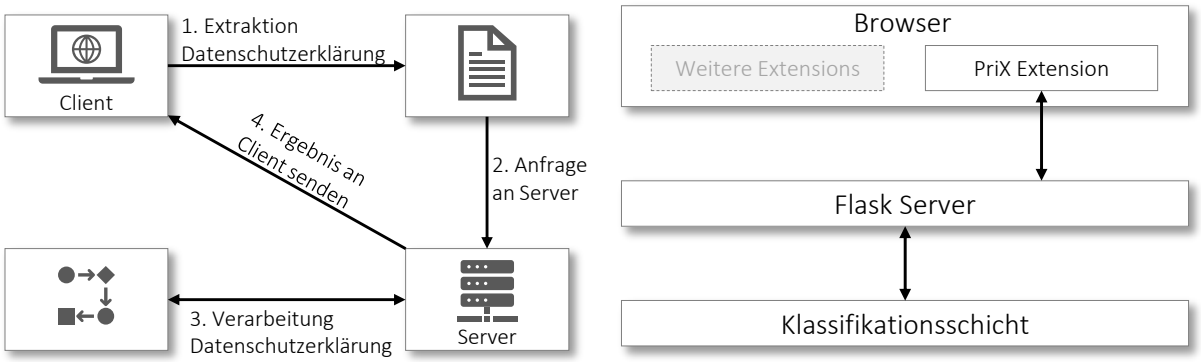
Eine *Visual Explanation* ist ein 3-Tupel $VE := (X, A, V_x)$, wobei X als Aspekt die Datenschutzerklärung darstellt, der Adressat A der Endbenutzer ist, der die Erklärung erhalten soll und V_x eine visuelle Repräsentation einer Untermenge von X ist.

Eine visuelle Repräsentation V_x besteht aus zwei Komponenten. Die erste Komponente von V_x ist eine Datenkategorie, die einen Abschnitt der DSE beschreibt, z.B. „Third Party Collection“, wenn eine Internetseite die Daten mit anderen Parteien teilt. Diese Kategorie fasst eine Textpassage zusammen und gibt relevante Informationen zum Verständnis sowie zur Unterstützung der Entscheidungsfindung des Nutzers, was wiederum als Erklärung betrachtet werden kann. Um die Erklärung zu bereichern und um ihr Verständnis zu erhöhen, sind Icons die zweite Komponente. Icons sind ein wesentlicher Bestandteil der Gestaltung von Benutzeroberflächen. Sie machen eine Benutzeroberfläche benutzbar, helfen dem Benutzer, sie zu verstehen und beschleunigen sogar den Verständnisprozess [308, 309, 310, 311]. Wir haben Icons des *Verein PRIVACY ICONS*[5] verwendet [312]. Jede Kategorie erhält hierbei ihr entsprechendes Icon (siehe Abschnitt C.1).

[5] `https://privacy-icons.ch`

5.1.2 Implementation - Architektur und Technologie

Für die Implementation von PriX wurde eine Client-Server-Architektur gewählt, die auf drei Schichten aufbaut. Eine Übersicht hierzu ist in Abbildung 5.3 dargestellt. Wie eingangs beschrieben ist PriX als Browser-Erweiterung konzipiert, basiert auf der WebExtension API[6] und ist somit kompatibel mit Chromium-basierten Browsern wie Google Chrome, Microsoft Edge, Opera, Vivaldi und auch Mozilla Firefox.



(a) Client-Server-Aufbau **(b)** Schichten-Aufbau

Abbildung 5.3: Architekturübersicht PriX

Der Server verwendet Flask[7] als Web-Framework und ist in Python implementiert, ebenso wie das Analysemodul (Klassifikationsschicht in Abbildung 5.3b). Die Kommunikation zwischen Client und Server findet über eine REST-Schnittstelle[8] statt. Hat PriX eine DSE identifiziert, sendet es diese automatisch an den Server (siehe Abbildung 5.3a). Der Server nimmt die DSE entgegen und reicht diese weiter an das Analysemodul. Da die DSE als HTML[9]-Code vorliegt, wird diese zunächst geparst und anschließend über die in Absatz 5.1.1.2 beschriebene Klassifizierungs-Pipeline verarbeitet. Das Ergebnis wird dann zurück an den Client gesendet und dort visuell aufbereitet.

5.1.2.1 UI Design von PriX

Die Benutzeroberfläche von PriX ist in Abbildung 5.4 dargestellt. Die grün eingekreisten Zahlen sowie der blau umrandete Kasten dienen hier ausschließlich Erläuterungszwecken. Letzterer markiert die beiden Komponenten einer *Visual Explanation*, bestehend aus dem Privacy Icon und der jeweiligen Datenkategorie. ① gibt den Domänennamen an, von der die

[6] `https://developer.chrome.com/docs/extensions/reference/`, Letzter Zugriff: 09.03.2023
[7] `https://flask.palletsprojects.com`, Letzter Zugriff: 09.03.2023
[8] `https://tinyurl.com/ymtffyyn`, Letzter Zugriff: 09.03.2023
[9] HTML bedeutet Hypertext Markup Language, `https://www.w3schools.com/html/`

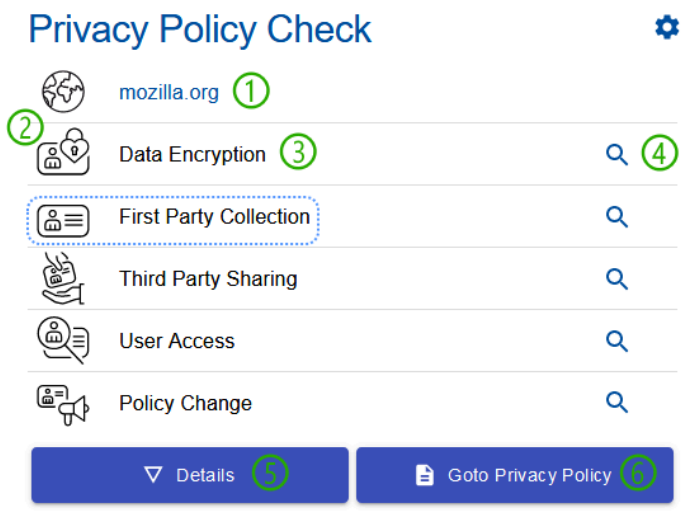


Abbildung 5.4: Benutzeroberfläche von PriX, annotiert mit Zahlen zur Erläuterung

DSE stammt. (2) ist das Privacy Icon der jeweiligen Kategorie. (3) gibt den Namen der jeweiligen Datenkategorie an. (4) Beim Klick auf die Lupe gelangt der Nutzer direkt zu dem entsprechenden Abschnitt in der DSE. (5) Ein Klick auf die Schaltfläche „Details“ zeigt Details (siehe Abbildung C.2) zur aktuellen DSE an. (6) Ein Klick auf „Goto Privacy Policy“ öffnet die DSE in einem neuen Browser-Fenster.

5.1.3 Evaluation durch Benutzerstudie

Um das Ziel unserer Studie zu formulieren, haben wir das Goal Definition Template [53, 54] angewandt. Die Formulierung lautet wir folgt:

> ***Goal Definition:*** **Analysiere** das Konzept von PriX **zum Zweck der** Evaluierung von Auswirkungen und Wahrnehmung der Software-seitigen Unterstützung für DSEs **in Bezug auf** Benutzerfreundlichkeit, Effektivität und Verständlichkeit **aus der Sicht von** Endnutzern **im Kontext** einer Online-Studie unter Verwendung der Think-Aloud-Methode.

Daraus ergeben sich die folgenden in Abbildung 5.5 abgebildeten Fragestellungen, die den Kern unserer Evaluation bildeten mit samt den damit verbundenen Metriken.

(1) **Auffinden der DSE (Q1):** Da es keinen einheitlichen Standard gibt, wie eine DSE auf einer Internetseite eingebettet sein muss, untersuchten wir beim ersten Aspekt, ob PriX Endbenutzer dabei unterstützt, die DSE einer Internetseite effizienter und vor allem schneller zu

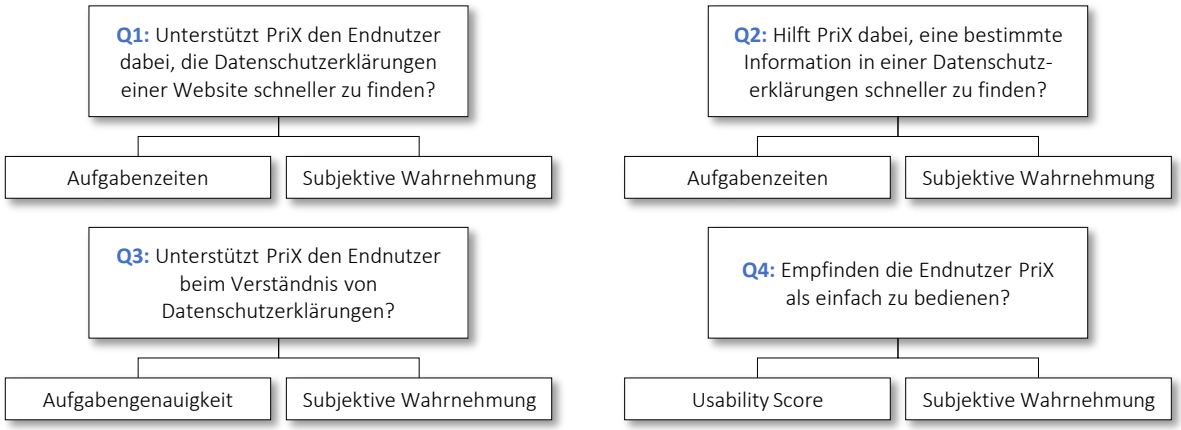


Abbildung 5.5: Fragen und damit verbundene Metriken

finden.

② **Auffinden bestimmter Informationen in der DSE (Q2):** In Fällen, in denen Benutzer nur bestimmte Aspekte der Datenverarbeitung kennen lernen wollen und nicht die Zeit haben, die gesamte DSE zu lesen, kann das Auffinden einer bestimmten Information eine Herausforderung darstellen. Viele DSEs bieten kein Inhaltsverzeichnis, mit dem die Benutzer direkt zu dem Abschnitt navigieren könnten, der sie interessiert. Ein Inhaltsverzeichnis könnte den Benutzern helfen, bestimmte Informationen schneller zu finden. Daher ist der zweite Aspekt, zu untersuchen, ob unser Tool den Endnutzern hilft, eine bestimmte Information schneller zu finden.

③ **Verstehen einer DSE (Q3):** Ohne einen juristischen Hintergrund sind die Seiten oft schwer zu verstehen und die Informationen sind in langen Textpassagen versteckt. Dies kann die Nutzer überfordern und dazu führen, dass die Seiten gar nicht gelesen werden. Aus diesem Grund analysieren wir, ob PriX dem Endbenutzer hilft, eine DSE zu verstehen.

④ **Wahrgenommene Benutzerfreundlichkeit (Q4):** Die Benutzerfreundlichkeit von Software ist ein wichtiger Punkt für die Akzeptanz und Nutzung eines Systems. Schlechte Usability kann weitreichende Folgen haben, die von der Verärgerung der Benutzer bis zur Gefährdung ihres Lebens reichen [313]. Daher untersuchen wir auch die Usability von PriX und die Akzeptanz des Tools und identifizieren mögliche Usability-Nachteile.

5.1.3.1 Design der Studie

Um das oben genannte Ziel zu erreichen und die Aspekte zu analysieren, haben wir einen *Synchronous Remote Usability Test* [314] mit einem Online-Fragebogen kombiniert. Aufgrund

der Pandemie war es nicht möglich, eine Usability-Studie in Präsenz durchzuführen. Brush et al. [315] stellten jedoch fest, dass es keinen quantitativen Unterschied zwischen einer Remote- und einer lokalen Studie gibt. Um die Qualität unserer Umfrage zu gewährleisten, haben wir uns an etablierte Richtlinien für die Umfrage- und Usability-Testgestaltung gehalten [316, 317, 318]. Wir haben uns für ein Studien-Design mit zwei Gruppen entschieden, da wir Gewöhnungs- und Lerneffekt vermeiden wollten, die die Studienergebnisse beeinflussen könnten. Daher wurden die Teilnehmer nach dem Zufallsprinzip in zwei Gruppen aufgeteilt. Die *Experimentalgruppe*, die PriX während der gestellten Aufgaben einsetzte und die *Kontrollgruppe*, die die Aufgaben ohne die Verwendung von PriX zu bewältigen hatte. Beiden Gruppen wurden dieselben DSEs gezeigt, da dies sonst zu einer Verzerrung hätte führen können.

Beide Gruppen erhielten die gleichen Aufgaben, die auf zwei echten DSEs basierten. Eine stammte von *mozilla.org* und die andere von *netflix.com*. Wir haben diese DSEs ausgewählt, da die DSE von Mozilla im Verhältnis zur relativen Länge von DSEs recht kurz ist und die Mozilla Foundation nur wenige Informationen über die Nutzer sammelt. Im Gegensatz dazu ist die DSE von Netflix sehr umfangreich und enthält eine Reihe an Informationen über die Datenpraktiken des Anbieters. Die Experimentalgruppe erhielt keine Einführung in die Nutzung von PriX, sondern erfuhr zu Beginn des Experiments nur, über welches Symbol in der Symbolleiste des Browsers PriX geöffnet werden kann. Ein Überblick über die gestellten Aufgaben t_1 bis t_6 ist in Tabelle 5.3 zu finden.

Tabelle 5.3: Übersicht der Aufgaben t_1 bis t_6

Aufgaben-ID	Beschreibung
Mozilla	
t_1	Bitte suchen Sie nun die Datenschutzerklärung von mozilla.org.
t_3	Sammelt Mozilla Daten über Sie und wenn ja, auf welche Weise?
t_4	Beschützt Mozilla Ihre Daten und wenn ja, auf welche Weise?
Netflix	
t_2	Bitte suchen Sie nun die Datenschutzerklärung von netflix.com.
t_5	Nutzt Netflix Cookies? Wenn ja, wieso?
t_6	Sammelt Netflix Daten über Sie und wenn ja, auf welche Weise?

5.1.3.2 Durchführung der Studie

Da es sich um eine Online-Studie handelt, haben wir Software wie TeamViewer[10] bzw. AnyDesk[11] verwendet, um den Teilnehmern Zugang zu unserem Studien-Computer zu gewähren. Auf diese Weise brauchten die Teilnehmer keine Software oder Plugins, auf ihren eigenen

[10] `https://www.teamviewer.com`
[11] `https://anydesk.com`

Computern zu installieren. Darüber hinaus haben wir BigBlueButton[12] für die Kommunikation genutzt. Wir nutzten OBS[13], um den Bildschirm des Computers des Experimentators aufzuzeichnen und das Experiment anschließend zu analysieren. Diese Bildschirmaufzeichnung ermöglichte es uns auch, die Aufgabenzeiten genau zu messen. Während der Studie wurde zudem die *Think-Aloud*-Methode [319] angewandt, deren Daten ebenfalls in die anschließende Analyse einflossen.

5.1.3.3 Datenerhebung und Analyse

Wir haben die Studie zwischen April und Mai 2021 durchgeführt. Die Studie wurde mit insgesamt 65 Teilnehmern durchgeführt. Die Experimentalgruppe bestand aus 33 Teilnehmern, während die Kontrollgruppe 32 Teilnehmer umfasste. Die Teilnehmer wurden den jeweiligen Gruppen randomisiert zugewiesen. Die Teilnehmerakquise fand über akademische Kontakte und Anfragen an Studierende der Leibniz Universität Hannover statt.

Variablen Während unserer Studie sammelten wir Daten zu den entsprechenden Variablen, die bei der Bewertung der vier genannten Aspekte halfen (siehe Abbildung 5.5). Insbesondere haben wir die für die Lösung bestimmter Aufgaben benötigten Zeiten, die Genauigkeit der Aufgaben und die Benutzerfreundlichkeitsbewertung gemessen. Diese Metriken wurden wie in Tabelle 5.4 angegeben definiert und berechnet. Zusätzlich zu diesen quantifizierbaren Me-

Tabelle 5.4: Metriken - Definition und wie sie berechnet werden

Variable	Beschreibung
Aufgabenzeit	(in Sekunden) ist definiert als die Zeit, die benötigt wird, um eine bestimmte Aufgabe zu lösen, z.B. das Auffinden der DSE oder einer entsprechenden Information darin. Es wird nicht zwischen richtiger und falscher Lösung differenziert.
Aufgabengenauigkeit	ist definiert als eine Ordinalskala, die von *A* bis *C* reicht. *A* bedeutet, der Teilnehmer konnte die Aufgabe korrekt lösen. *B* bedeutet, der Befragte befand sich im richtigen Abschnitt der DSE, konnte aber letztendlich die geforderten Informationen nicht korrekt angeben. *C* bedeutet, dass die Aufgabe nicht beantwortet werden konnte oder dass die gegebene Antwort falsch war.
Usability Score	ist definiert als der SUS [320], der als eine etablierte Metrik zur Bewertung der Software-Usability gilt [321].

triken haben wir zusätzlich die subjektiven Wahrnehmungen der Teilnehmer berücksichtigt. Die geschah über Likert-Skalen für die Fragen Q1-Q3.

[12] `https://bigbluebutton.org`
[13] `https://obsproject.com`

Qualitative Analyse In den Fragebögen wurden offene und geschlossene Fragen gestellt. So wurden beispielsweise offene Fragen zur Aufgabe t_6 gestellt sowie Fragen für die Experimentalgruppe zu den Visual Explanations (z.B. was eine bestimmte Datenkategorie bedeuten könnte). Für die Analyse der offenen Fragen wurde ein offener Kodierungsansatz [230], analog zu dem in Abschnitt 4.1.2.2 beschriebenem verwendet. Wir verwendeten zwei aufeinanderfolgende Kodierungszyklen. Wir nutzten *In Vivo Coding* [231] während des ersten Kodierzyklus, um die Perspektiven der Teilnehmer zu reflektieren. Der zweite Kodierzyklus besteht aus dem *Pattern Coding* [232], um die anfänglichen Codes in eine geringere Anzahl von Themen oder Konstrukten zu unterteilen. Der Kodierungsprozess wurde unabhängig voneinander (von Lukas Köhler und mir) durchgeführt. Bei Unstimmigkeiten diskutierten die Autoren diese, bis sie einen gemeinsamen Konsens gefunden hatten.

Quantitative Analyse Um den Einfluss von PriX in Bezug auf die vier zuvor genannten Aspekte zu messen, haben wir die Daten beider Studiengruppen auf Unterschiede hin analysiert. Insbesondere haben wir die in Tabelle 5.5 zusammengefassten Nullhypothesen getestet.

Tabelle 5.5: Nullhypothesen zum Testen der Daten auf Unterschiede zwischen den beiden Gruppen

ID	**Hypothesen** *Es gibt keinen Unterschied zwischen den beiden Gruppen...*
$H1_0$	hinsichtlich der Zeit, die benötigt wird, um die DSE zu finden.
$H1.1_0$	in Bezug auf die Zeit, die benötigt wird, um die DSE von Mozilla zu finden.
$H1.2_0$	hinsichtlich der Zeit, die benötigt wird, um die DSE von Netflix zu finden.
$H2_0$	in Bezug auf die Zeit, die benötigt wird, um eine bestimmte Information zu finden.
$H2.1_0$	in Bezug auf die Zeit, die benötigt wird, um Informationen über die Datenerhebung in der DSE von Mozilla zu finden.
$H2.2_0$	in Bezug auf die Zeit, die benötigt wird, um Informationen zum Datenschutz in der DSE von Mozilla zu finden.
$H2.3_0$	in Bezug auf die Zeit, die benötigt wird, um Informationen über die Verwendung von Cookies in der DSE von Netflix zu finden.
$H2.4_0$	in Bezug auf die Zeit, die benötigt wird, um Informationen über die Datenerfassung in der DSE von Netflix zu finden.

Die Hypothesen H1 und H2 beziehen sich auf die ersten beiden Aspekte (d. h. das Auffinden der DSE und das Auffinden einer bestimmten Information darin).

Da die Aufgabenzeiten positiv verzerrt sind, haben wir eine logarithmische Transformation auf die Rohzeiten angewandt [322], um die Genauigkeit der Ergebnisse zu verbessern [323]. Nach dem zentralen Grenzwertsatz kann bei einer ausreichenden Stichprobengröße ($n > 30$) eine Normalverteilung angenommen werden [324]. Daher haben wir den z-Test verwendet, um $H1.1_0$ und $H1.2_0$ zu analysieren. Da wir jedoch zwei Tests durchgeführt haben, um die

Haupthypothese H1 zu testen, wenden wir die Bonferroni-Korrektur an, was zu einem bereinigten p-Wert von $p_{corr} = p/2 = 0,025$ führt. Das heißt, sobald eine der beiden Hypothesen H1.1 und H1.2 mit einem p-Wert kleiner als p_{corr} signifikant ist, können wir davon ausgehen, dass es Unterschiede zwischen den beiden Gruppen gibt, und somit können wir H1 verwerfen.

Um die Zeiten zu vergleichen, die benötigt wurden, um eine bestimmte Information in der DSE zu finden, haben wir ein ähnliches Verfahren angewandt: Beide Gruppen bekamen die gleichen Aufgaben (t_3 - t_6). Die Aufgaben t_3 und t_4 bezogen sich auf Mozillas DSE und t_5 und t_6 auf die DSE von Netflix. Nach dem zentralen Grenzwertsatz kann auch hier eine Normalverteilung angenommen werden, da eine ausreichende Stichprobengröße ($n > 30$) vorliegt. Wir haben auch hier eine logarithmische Transformation auf die Aufgabenzeiten angewendet und den z-Test verwendet, um $H2.1_0$ bis $H2.4_0$ auf statistische Signifikanz zu testen. Da wir wiederum vier Tests zur Überprüfung der Haupthypothese H2 durchgeführt haben, wenden wir die Bonferroni-Korrektur an, die zu einem bereinigten p-Wert von $p_{corr} = p/4 = 0,0125$ führt. Das heißt, sobald einer der beiden Tests mit einem p-Wert kleiner als p_{corr} signifikant ist, können wir davon ausgehen, dass es Unterschiede zwischen den beiden Gruppen gibt.

Für den dritten Aspekt, das Verständnis einer DSE, analysierten wir die Genauigkeit der Ergebnisse der Teilnehmer bei der Durchführung einer bestimmten Aufgabe unter Verwendung des in Absatz 5.1.3.3 beschriebenen Schemas. Für den vierten Aspekt haben wir für die Bewertung der Benutzerfreundlichkeit den SUS herangezogen.

5.2 Ergebnisse der Benutzerstudie

Die in Abschnitt 5.1.3 beschriebenen Benutzerstudie wurde mit insgesamt 65 Teilnehmern durchgeführt, wovon 33 Probanden der Experimentalgruppe und 32 Probanden der Kontrollgruppe zugewiesen wurden. Die Zuweisung erfolgte randomisiert. Voraussetzung für die Studienteilnahme waren neben rudimentären Computer-Kenntnissen lediglich grundlegende Englischkenntnisse, da die Benutzerstudie (die verwendeten DSEs) englische Inhalte verwendete. Mit rudimentären Computer-Kenntnissen ist gemeint, dass die Teilnehmer der Studie lediglich einen Browser bedienen können mussten.

5.2.1 Demographische Daten und Privatsphäreverhalten

Das Alter der Teilnehmer reichte von 18 bis 59 Jahren (M=23,8, SD=5,4). Sofern nicht anders angegeben stehen die Abkürzungen *M* für das arithmetische Mittel und *SD* für die Standardabweichung. Die Mehrheit der Teilnehmer (80%) waren Studenten. Insgesamt 64 Teilnehmer (98,46%) gaben an, dass sie das Internet mindestens einmal am Tag nutzen. Der Aussage *„Privatsphäre ist wichtig für mich“* stimmten 49 Probanden (75,38%) zu. Dies steht im Kontrast zu

den Aussagen der Probanden, wie oft sie darauf achten, ob eine Internetseite eine DSE hat. In diesem Fall gaben 72,31% der Befragten an, dass sie dies selten bis nie tun. Darüber hinaus gaben nur sechs Teilnehmer (9,23%) an, dass sie die DSE der Internetseite, die sie gerade besuchen, häufig lesen. Drei Teilnehmer (4,62%) sagten, dass sie sie immer lesen, und neun (13,84%) gaben an, dass sie dies gelegentlich tun. Darüber hinaus gaben 50 (76.92%) Teilnehmer an, dass sie bei unterschiedlichen Diensten bzw. Anbietern auch unterschiedlichen Wert auf den Schutz ihrer Privatsphäre legen. Als Grund dafür gaben die Teilnehmer an, dass die Art und Menge der Daten für sie relevant sei, die der jeweilige Dienst über sie sammelt. Bei Anbietern, die mit Hilfe ihrer Internetseiten viele persönliche Daten über die Teilnehmer sammeln, gaben diese an, auch mehr Wert auf ihren Datenschutz zu legen.

5.2.2 Auffinden einer Datenschutzerklärung

Zunächst haben wir die Zeit verglichen, die benötigt wurde, um eine DSE mit oder ohne PriX zu finden. Die Probanden beider Gruppen wurden gebeten, die vorgegebene Internetseite (t_1 : mozilla.org, t_2 : netflix.com) zu öffnen und dann die jeweilige DSE zu suchen und zu öffnen. Um die Hypothese H1 zu testen (siehe Tabelle 5.5), haben wir die Hypothesen H1.1 und H1.2 mit dem z-Test getestet.

Tabelle 5.6: Ergebnisse des z-Tests, um H1 zu testen

ID	Ergebnisse	Signifikant?
$H1_0$		Ja
$H1.1_0$	$z \approx -7.59,\ p < 0.01$	Ja
$H1.2_0$	$z \approx -2.13,\ p \approx 0.03$	Ja

Die Ergebnisse(siehe Tabelle 5.6) zeigen einen signifikanten Unterschied zwischen den beiden Gruppen zugunsten der Experimentalgruppe, so dass wir die Nullhypothesen $H1.1_0$ und $H1.2_0$ verwerfen können. Da insbesondere der p-Wert für die DSE von Mozilla unter dem korrigierten p-Wert von $0,025$ liegt, können wir auch die Haupthypothese $H1_0$ verwerfen und davon ausgehen, dass die Verwendung von PriX die Zeit zum Auffinden einer DSE signifikant reduziert.

Auf die Frage nach dem Grad der Zustimmung zu der Aussage *„Die DSE war leicht zu finden“* stimmten alle Teilnehmer der Experimentalgruppe für beide DSEs (Mozilla und Netflix) zu (oder stimmten voll zu). In der Kontrollgruppe stimmten 12,49% voll zu, 37,5% stimmten zu, 18,75% stimmten weder zu noch lehnten sie ab, 21,88% stimmten nicht zu und 9,38% lehnten die Aussage für die DSE von Mozilla absolut ab. Im Falle von Netflix stimmten 53,11% voll zu, 37,5% stimmten zu und jeweils 3,13% stimmten weder zu noch stimmten sie zu, stimmten nicht

zu oder stimmten absolut nicht zu (siehe Tabelle 5.7). Die Werte zeigen für die DSE von Mozilla einen höheren Wert bei der Nicht-Zustimmung. Eine möglicher Erklärung ist, da Mozilla für jedes ihrer Produkte (Firefox, Thunderbird, etc.) eine eigene DSE anbietet, jedoch in der Aufgabe nach einer spezifischen DSE gefragt wurde. Hingegen verlinkt Netflix nur eine DSE auf ihrer Internetseite.

Tabelle 5.7: Ergebnisse Kontrollgruppe - DSE war einfach zu finden

	Stimme voll zu	**Stimme zu**	**Neutral**	**Stimme nicht zu**	**Stimme absolut nicht zu**
DSE Mozilla	12,49%	37,5%	18,75%	21,88%	9,38%
DSE Netflix	53,11%	37,5%	3,13%	3,13%z	3,13%

Eine weitere Analyse der aufgezeichneten Daten zeigte auch, dass einige Probanden aus der Kontrollgruppe, obwohl sie während unseres Experiments Probleme bei der Suche nach der DSE hatten (sowohl bei Mozilla als auch bei Netflix), später behaupteten, keine Probleme gehabt zu haben. Ein Grund dafür könnte sein, dass es den Teilnehmern unangenehm war, ihre möglichen Probleme zuzugeben oder dieses Problem ihnen möglicherweise auch gar nicht bewusst war. Zusammenfassend können wir sagen, dass der z-Test signifikante Unterschiede in den Aufgabenzeiten nachgewiesen hat und der Vergleich der Antworten der Teilnehmer auch bestätigt, dass PriX die Endnutzer dabei unterstützen kann, eine DSE einer Internetseite schneller zu finden.

5.2.3 Finden spezifischer Informationen in einer Datenschutzerklärung

Nachdem die Probanden im vorangegangenen Schritt eine DSE finden mussten, wurden Sie nun gebeten, bestimmte Informationen einer DSE zu entnehmen. Hierbei wurden ebenfalls wieder die Zeiten verglichen, die die Probanden benötigten, um ihre Aufgaben abzuschließen. Beide Gruppen bekamen die gleichen Aufgaben (t_3 - t_6). Die Beschreibung der Aufgaben ist in Tabelle 5.3 zu finden. Um die Haupthypothese H2 zu testen, haben wir die Hypothesen $H2.1_0$ bis $H2.4_0$ ebenfalls mit dem z-Test überprüft. Die Ergebnisse sind in Tabelle 5.8 zusammengefasst.

Bei den Aufgaben, die sich auf Mozillas DSE beziehen, wurde ein signifikanter Unterschied bei den Aufgabenzeiten zu Gunsten der Experimentalgruppe festgestellt. Die Verwendung PriX verringert hier also signifikant die Zeit, die benötigt wird, um eine bestimmte Information in der DSE von Mozilla zu finden. Im Gegensatz dazu wurde kein signifikanter Unterschied

Tabelle 5.8: Ergebnisse des z-Tests, um H2 zu testen

ID	Ergebnisse	Signifikant?
$H2_0$		Ja
$H2.1_0$	$z \approx -2.10, p \approx 0.04$	Ja
$H2.2_0$	$z \approx -3.86, p < 0.01$	Ja
$H2.3_0$	$z \approx -0.26, p \approx 0.80$	Nein
$H2.4_0$	$z \approx -1.19, p \approx 0.24$	Nein

in Bezug auf die DSE von Netflix festgestellt. Auch wenn hier kein signifikanter Unterschied nachgewiesen werden konnten, war die Experimentalgruppe auch hier im Durchschnitt schneller (siehe Tabelle 5.9).

Tabelle 5.9: Werte der Aufgabenzeiten - Frage 2

Aufgabe	**M** (Experimentalgruppe)	**M** (Kontrollgruppe)	**SD** (Experimentalgruppe)	**SD** (Kontrollgruppe)
t_5	138,27	145,38	81,69	85,55
t_6	54,09	67,69	60,36	51,99

Da der p-Wert für H2.2 insgesamt unter dem korrigierten p-Wert von 0,0125 liegt, können wir H2 dennoch verwerfen und somit davon ausgehen, dass es einen signifikanten Unterschied gibt in der benötigten Zeit bei der Suche nach einer bestimmten Information in einer DSE.

Zu beobachten war bei der Kontrollgruppe, dass nur etwa die Hälfte der Probanden die Suchfunktion des Browsers nutzte, um die spezifischen Informationen zu finden. Wir haben die Probanden im Vorfeld weder dazu ermuntert dies zu tun, noch ihnen das Nutzen der Suchfunktion untersagt. Nach Abschluss der Aufgaben fragten wir die Teilnehmer, ob die Informationen zu den Aufgaben leicht zu finden waren. Für Mozillas DSE stimmten 72,73% der Teilnehmer der Experimentalgruppe voll zu, 24,24% stimmten zu, und ein Teilnehmer stimmte weder zu noch widersprach er (vgl. Tabelle 5.10). 46,88% der Probanden der Kontrollgruppe stimmten zu, 40,62% stimmten voll zu und 12,5% stimmten weder zu noch widersprachen sie (vgl. Tabelle 5.11).

Tabelle 5.10: Ergebnisse Experimentalgruppe - Informationen waren einfach zu finden

	Stimme voll zu	**Stimme zu**	**Neutral**	**Stimme nicht zu**	**Stimme absolut nicht zu**
DSE Mozilla	72,73%	24,24%	3,03%	-	-
DSE Netflix	42,43%	48,48%	6,06%	-	-

Auf die Frage nach Netflix' DSE stimmten 42,43% der Teilnehmer der Experimentalgruppe voll zu, 48,48% stimmten zu, 6,06% stimmten weder zu noch widersprachen sie, und ein Proband stimmte nicht zu (vgl. Tabelle 5.11). In der Kontrollgruppe hingegen stimmten 18,75% nicht

zu, 31,25% stimmten weder zu noch nicht zu, 21,87% stimmten zu und 28,13% stimmten voll zu (vgl. Tabelle 5.11).

Tabelle 5.11: Ergebnisse Kontrollgruppe - Informationen waren einfach zu finden

	Stimme voll zu	**Stimme zu**	**Neutral**	**Stimme nicht zu**	**Stimme absolut nicht zu**
DSE Mozilla	40,62%	48,88%	12,50%	-	-
DSE Netflix	28,13%	21,87%	31,25%	18,75%	-

PriX hebt die spezifische Textpassage in einer DSE hervor, die der jeweiligen Datenkategorie entspricht. 93,9% der Teilnehmer stimmten zu, dass diese Funktion sehr nützlich ist, um bestimmte Informationen in einer DSE zu finden. Es hat sich auch gezeigt, dass die visuellen Erklärungen von PriX bei umfangreicheren DSEs einen Geschwindigkeitsvorteil bieten, der in diesem Fall zwar nicht signifikant, aber im Durchschnitt vorhanden war.

5.2.4 Verständnis einer Datenschutzerklärung

Im nächsten Schritt haben wir analysiert, ob die PriX das Verständnis einer DSE unterstützt. In Tabelle 5.12 ist eine Übersicht über die Genauigkeit der Ergebnisse für die Aufgaben t_3 - t_6 dargestellt. Für eine bessere Übersicht sind die Ergebnisse aggregiert.

Tabelle 5.12: Genauigkeit der Ergebnisse - Frage 3

	Experimentalgruppe			**Kontrollgruppe**		
	A	**B**	**C**	**A**	**B**	**C**
$t_{3,4}$	87,8%	4,6%	7,6%	70,31%	18,75%	10,94%
$t_{5,6}$	65,2%	18,1%	16,7%	62,4%	14,1%	23,5%

Ein Chi-Quadrat-Test auf Unabhängigkeit wurde durchgeführt, um die Beziehung zwischen den beiden Gruppen und die Genauigkeit der Ergebnisse zu untersuchen. Der Zusammenhang zwischen diesen Variablen war signifikant, χ^2 (2, N = 65) = 7,3451, p = 0,0029. Die Probanden in der Experimentalgruppe lösten die Aufgaben mit größerer Wahrscheinlichkeit richtig.

Nachdem die Teilnehmer die Aufgaben gelöst hatten, fragten wir sie, ob sie wüssten, welche Daten der Anbieter sammelt und was mit ihnen geschieht. Die Probanden der Experimentalgruppe gaben an, etwas besser über die Datenpraktiken des Anbieters informiert zu sein. Hier antworteten 31,28% mit ja, 68,13% mit nein. In der Kontrollgruppe antworteten 21,87% mit ja und 78,13% mit nein.

Betrachtet man die Daten in Tabelle 5.12 des signifikanten Ergebnisses des Chi-Quadrat-Tests und die Tatsache, dass mehr Teilnehmer in der Experimentalgruppe angaben, besser

über die Datenpraktiken des Anbieters informiert zu sein, könnte eine Schlussfolgerung gerechtfertigt sein, dass PriX nicht nur das Verständnis verbessert, sondern dem Nutzer auch ein größeres Gefühl des Informiert-Seins vermitteln kann. Allerdings können wir diese Schlussfolgerung nicht allein auf der Grundlage der vorgelegten Daten ziehen, da sie auch auf die unterschiedlichen Standpunkte der Befragten zurückzuführen sein könnte. Aufgrund unserer Zufallszuweisung könnte eine größere Anzahl misstrauischerer Nutzer in der Kontrollgruppe gelandet sein.

Die Fragen zum Verständnis der jeweiligen DSEs beantworteten die Teilnehmer der Kontrollgruppe wie in Tabelle 5.13 zu sehen. Der Experimentalgruppe wurde diese spezielle Frage zum subjektiven Eindruck des Verständnisses nicht gestellt, da sie sich hauptsächlich mit einer durch PriX aufbereiteten DSE beschäftigten und wir hier wissen wollten, wie die Teilnehmer die visuellen Erklärungen wahrnahmen.

Tabelle 5.13: Ergebnisse Kontrollgruppe - DSE war einfach zu verstehen

	Stimme voll zu	**Stimme zu**	**Neutral**	**Stimme nicht zu**	**Stimme absolut nicht zu**
DSE Mozilla	43,75%	43,75%	9,38%	3,13%	-
DSE Netflix	28,13%	31,25%	31,25%	9,38%	-

Die Experimentalgruppe empfand die Datenkategorien als unterstützend für ihr Verständnis, was durch eine aggregierte Zustimmungsquote von 87,88% belegt wird. 6,1% stimmten nicht zu, dass die Datenkategorien sie beim besseren Verständnis der DSE unterstützen. Die Symbole wurden von 64,52% der Befragten als hilfreich empfunden, während 9,68% sie nicht als hilfreich empfanden. Zwei Probanden haben sie gar nicht wahrgenommen. Wir gehen davon aus, dass diese Wirkung darauf zurückzuführen ist, dass die Benutzer die Symbole zum ersten Mal sehen. Bei häufigerem Gebrauch könnten die Benutzer in der Lage sein, das Symbol zu erkennen und mit der entsprechenden Datenkategorie in Verbindung zu bringen. Einige Kommentare der Teilnehmer unterstützen diese Hypothese. PriX gibt auch detaillierte Textinformationen über eine bestimmte Datenkategorie. 78,8% der Teilnehmer empfanden diese Informationen als nützlich, 18,17% stimmten weder zu noch widersprachen sie, und nur ein Teilnehmer (3,03%) empfand sie als nicht nützlich. Wir konnten auch nachweisen, dass PriX den Probanden half, effizienter mit DSEs umzugehen. Außerdem hilft es den Nutzern, eine DSE besser zu verstehen, was durch die von uns analysierten objektiven Metriken (Genauigkeit der Aufgabenergebnisse) und die subjektiven Ergebnisse hinsichtlich der Wahrnehmung der Tool-Unterstützung durch die Nutzer bestätigt wird.

5.2.5 Wahrgenommene Usability

Um die empfundene Usability der Teilnehmer in der Experimentalgruppe zu untersuchen haben wir den SUS [320] für PriX berechnet. Beim SUS handelt es sich um eine „Zehn-Punkte-Skala, die einen ganzheitlichen Überblick über die subjektive Bewertungen der Benutzerfreundlichkeit“ [320] einer Software gibt. Die berechnete SUS-Punktzahl für PriX lag im Durchschnitt bei 87,95, was 19,75 Punkte höher ist als die einer durchschnittlichen Webanwendung [325]. Das berechnete 95% Konfidenzintervall $[84,49,\ 91,41]$ zeigt, dass PriX ein System mit mindestens *überdurchschnittlicher* oder *fast ausgezeichneter* Benutzerfreundlichkeit ist [325]. 96,7% der Teilnehmer der Experimentalgruppe gaben an, dass PriX es ihnen erleichtert hat, einen Überblick über die DSEs zu gewinnen. Nur ein Teilnehmer stimmte weder zu noch stimmte er nicht zu. Wie die Ergebnisse zeigen, bietet PriX eine wirksame Unterstützung im Umgang mit DSEs. Die Probanden empfanden die visuellen Erklärungen als hilfreich, sowohl für das Auffinden von Informationen als auch für das Verständnis einer DSE.

5.3 Einschränkungen und Bedrohung der Validität

In diesem Abschnitt diskutiere ich die Einschränkungen und Bedrohung der Validität bezugnehmend auf Wohlin et al. [53].

Konstruktvalidität Die Antworten der Teilnehmer beruhten auf ihren Selbstauskünften. Einige Teilnehmer haben möglicherweise behauptet, eine DSE besser zu verstehen, weil sie sich in einer Testsituation befanden. Der Versuchsleiter war während des gesamten Versuchs anwesend, um die Aufgabe zu erklären und bei Fragen zu helfen. Auch dies könnte einen Einfluss auf die Ergebnisse gehabt haben. Wir haben versucht, dieser Gefahr zu begegnen, indem wir den Versuchsleiter angewiesen haben, so unauffällig wie möglich zu sein.

Aussagekraft der Schlussfolgerung Wir haben Signifikanztests durchgeführt, um die Signifikanz der Ergebnisse zu gewährleisten. Trotz der Tatsache, dass die Gesamtstichprobe 65 Teilnehmer umfasste, können einige unserer Ergebnisse aufgrund der geringen Stichprobengröße gefährdet sein. Eine Wiederholung der Studie mit einer größeren Stichprobengröße und einer heterogeneren Verteilung ist erforderlich, um unsere Ergebnisse zu bestätigen.

Interne Validität Die Ergebnisse des Experiments könnten durch die Qualität der Fragen beeinflusst worden sein. Wir haben versucht, diese Gefahr zu verringern, indem wir unsere Studie pilotiert haben; außerdem wurden die Fragen in der Muttersprache der Probanden verfasst, um Missverständnisse zu vermeiden.

Externe Validität Die Auswahl der Teilnehmer könnte die Gültigkeit unserer Ergebnisse gefährdet haben. 80% unserer Teilnehmer gaben an, Studenten zu sein. Obwohl für unsere Studie kein spezifisches Hintergrundwissen erforderlich war, spiegeln die Ergebnisse möglicherweise nicht die Bedürfnisse und Wahrnehmungen der gesamten Population wider. Dies stellt eine Einschränkung für die Verallgemeinerung der Ergebnisse dar.

5.4 Fazit

DSEs sind der wichtigste von Unternehmen genutzte Kanal, um die Endnutzer über die Datenerfassung und den damit verbundenen Datenpraktiken zu informieren. Leider versetzt die bloße Existenz einer DSE die Endnutzer nicht in die Lage, bewusste Entscheidungen über die Nutzung eines bestimmten Dienstes in Bezug auf die Privatsphäre zu treffen. Bekannte Gründe dafür wie zu lange und schwer zu verstehende Texte sind eingangs bereits diskutiert worden. Viele Nutzer vernachlässigen ihre Privatsphäre aufgrund mangelnder Kenntnisse über das Thema Datenschutz, mentaler Überforderung oder schlichtweg Zeitmangel [184, 109, 291]. Doch wenn DSEs weder gelesen noch verstanden werden und zudem die meist einzige Quelle an Informationen zu Datenpraktiken sind, auf welcher Grundlage können die Nutzer dann Entscheidungen über ihre Privatsphäre treffen? 75,38% der Probanden unserer Studie gaben an, dass ihnen ihre Privatsphäre wichtig ist. Auch andere Studien bestätigen das Interesse am Datenschutz bei einer Mehrheit der Nutzer [326, 327]. Dies unterstreicht zum einen die Notwendigkeit aber auch das Recht der Nutzer, informierte und bewusste Entscheidungen über ihre Privatsphäre treffen zu können, wie es beispielsweise in der DSGVO verankert ist.

Mit Bezug auf DSEs konnten wir aufzeigen, dass PriX diese einfacher zugänglich und verständlicher macht, indem z.B. *Visual Explanations* bereitgestellt werden und die Nutzer in die Lage versetzt werden, ihr Bewusstsein für die Privatsphäre zu verbessern. Unsere Studienergebnisse scheinen zu belegen, dass die gegebenen Erklärungen ein geeignetes Mittel sind, um das Verständnis der Nutzer für eine DSE zu fördern und zu verbessern. Unsere Ergebnisse zeigen auch, dass Endnutzer, die in der Lage sind, einfach auf Informationen zum Datenschutz zugreifen zu können, bereit sind, diese Informationen zu konsumieren. Dies wird durch die Tatsache untermauert, dass am Ende der Studie mehr als die Hälfte der Probanden aus der Experimentalgruppe gefragt hat, ob wir die PriX öffentlich verfügbar machen werden, da sie PriX gerne regelmäßig einsetzen würden. Dies deutet darauf hin, dass PriX ein geeigneter Vermittler im Umgang mit DSEs zu sein scheint und sich schlussendlich somit positiv auf das Datenschutzbewusstsein (Privacy Awareness) der Endnutzer auswirkt.

Es steht außer Frage, dass die heutigen Systeme uns in vielen Bereichen des alltäglichen Lebens unterstützen und helfen. Aber bedeutet der Zuwachs an Systemen, die immer komplexer und invasiver werden was unsere Privatsphäre angeht, gleichzeitig auch komplexere DSEs und somit immer ausgefeiltere Werkzeuge, die uns beim Umgang mit ihnen unterstützen? Das kann doch eigentlich nicht die Lösung für den Endbenutzer sein. Waldman [49] spricht hier genau das zugrunde liegende Problem an. Er sagt das Problem ist, dass DSEs „ohne Rücksicht auf die Bedürfnisse der realen Menschen erstellt werden. Sie werden von Anwälten und für Anwälte geschrieben und ignorieren die Art und Weise, wie die meisten von uns Entscheidungen über die Offenlegung von Informationen online treffen. Sie ignorieren auch die Einflüsse von Design, Ästhetik und Präsentation auf unsere Entscheidungsfindung“ [49]. Wichtig ist hier nun aber festzuhalten, dass DSEs nicht per se ein Problem sind. Sie haben absolut ihre Daseinsberechtigung und sind wichtig für den juristischen und regulatorischen Sektor. Sie sind aber **kein** Instrument, reale Menschen (Endbenutzer) über Datenpraktiken aufzuklären.

Offensichtlich muss hier nach anderen Lösungen geschaut werden, Transparenz bei Datenpraktiken zu schaffen und den Endbenutzer auf leicht verständliche, nachvollziehbare Art zu informieren und aufzuklären bei gleichzeitiger Wahrung und Einhaltung regulatorischer Vorgaben. In den folgenden Kapiteln meiner Dissertation stelle ich daher nun einen möglichen Lösungsansatz vor, wie man dieser Herausforderung benutzerzentriert begegnen könnte.

6

Erklärbarkeit und Privatsphäre - Eine nutzerzentrierte Lösung

Persönliche Daten sind zu einer wichtigen Währung im digitalen Zeitalter geworden [27, 28] und bei der Interaktion mit einem Informationssystem legen wir meist unbewusst eine ganze Reihe dieser persönlichen Daten über uns offen (siehe Abschnitt 1.1). Tatsächlich kann bereits das Design eines Systems eine essentielle Rolle bei der Wahrung und Einhaltung des Datenschutzes spielen [328]. Dieser Grundgedanke spiegelt sich auch im Konzept *Privacy by Design (PbD)* wider. Eine erste Arbeit, in der auf den Datenschutz eines Informationssystems aus der Perspektive eines (System-) Designers geschaut wird, wurde bereits 1995 veröffentlicht [329]. Der Terminus *Privacy by Design* selber tauchte aber erst im Jahre 2000 auf der *Conference on Computers, Freedom & Privacy*[1] auf. Maßgeblich geprägt wurde PbD als Entwicklungsansatz [9, 330], durch die kanadische Datenschützerin Ann Cavoukian[2]. Bei PbD handelt es sich um einen proaktiven Ansatz, bei dem der Datenschutz bereits von Beginn der Entwicklung an berücksichtigt wird, anstatt erst rückwirkend Datenschutzmaßnahmen zu ergreifen [190]. Also Prävention (Datenschutzrisiken im Vorfeld begegnen) statt Reaktion und somit möglicherweise kostspielige Anpassungen im Nachgang. Cavoukian sagt, dass PbD auf

[1] `http://cfp2000.org`, Letzter Zugriff: 04.04.2023

[2] Ann Cavoukian war als Informationsfreiheits- und Datenschutzbeauftragte in Ontario, Kanada tätig.

sieben Prinzipien basiert (siehe Tabelle 6.1), wobei im Laufe der Zeit auch das Prinzip der *Datensparsamkeit* Einzug gehalten hat [331, 332].

Im Zusammenhang mit der Datensparsamkeit sprechen Jiang et al. [333] vom **Prinzip der minimalen Asymmetrie** (englisch: *principle of minimum asymmetry* oder auch *minimal information asymmetry*), welches ein auf den Datenschutz ausgerichtetes Informationssystem berücksichtigen sollte. Die Autoren schreiben, „ein datenschutzfreundliches System sollte die Informationsasymmetrie zwischen Dateneigentümern, Datensammlern und Datennutzern minimieren“ [333]. Dies sollte durch die „*Verringerung* des Informationsflusses von Dateneigentümern zu Datensammlern und -nutzern“ und durch die „*Erhöhung* des Informationsflusses von Datensammlern und -nutzern zurück zu Dateneigentümern geschehen“ [333]. Ein **Dateneigentümer** (data owner) bezeichnet ein Individuum, dessen Daten verwendet werden oder auf die zugegriffen wird (z.B. Endnutzer), **Datensammler** (data collector) sind Individuen oder auch Systeme, die Informationen über Dateneigentümer sammeln, so Jiang et al. [333]. Weiter heißt es, dass **Datennutzer** (data user) Individuen oder Systeme sind, die diese Informationen nutzen und/oder verarbeiten.

Aktuell mangelt es allerdings noch an praktischem Wissen wie privatsphärefreundliche Informationssysteme gebaut werden können [334, 335] und auch an kollektiven Ressourcen (Wissen, Vorgehensweisen, etc.) darüber, was PbD in einer konkreten Anwendungsdomäne bedeutet [332]. Thiel et al. bestätigen dies und äußern, dass „zur Umsetzung in Unternehmen fehlen konkrete Anleitungen und Best-Practice-Sammlungen. Das führt dazu, dass PbD bisher von wenigen Akteur*innen eingesetzt wird“ [336]. Van Rest et al. geben zudem zu bedenken, dass es aktuell noch eine offene Frage ist, „was Bürger und Verbraucher tatsächlich über PbD wissen müssen, damit das Konzept ‚funktioniert‘“ [332], denn weiter heißt es „dass das Konzept des PbD derzeit nicht dazu geeignet ist, den […] Endnutzern Vertrauen (oder mangelndes Vertrauen) in bestimmte Systeme zu vermitteln“ [332], was aber primär auf die nicht erfolgte Umsetzung des Konzepts zurückzuführen ist bzw. dem Mangel an Wissen zum Thema Privatsphäre bei u.a. Entwicklern [337, 338, 339]. Denn im Grunde genommen soll PbD bei Endbenutzern für mehr Transparenz sorgen, Kontrolle über die eigenen Daten geben sowie für Datensparsamkeit sorgen und dem Prinzip der minimalen Asymmetrie folgen [336].

Eine Möglichkeit, die im Einklang mit dem PbD-Konzept und dem Prinzip der minimalen Asymmetrie steht, Endnutzer hinsichtlich ihrer Privatsphäre zu sensibilisieren und gleichzeitig über etwaige Datenpraktiken eines Anbieters zu informieren und aufzuklären, damit Endbenutzer bewusstere Entscheidungen treffen können, also ihre *explizite Zustimmung* (siehe Abschnitt 1.1) geben könnten, möchte ich im folgenden vorstellen.

Es existieren Studien, die den Einfluss von Erklärungen auf die User Experience untersuchen [340, 341] oder andere, die untersuchen welche Arten von Erklärungen in welchen

Tabelle 6.1: 7 Prinzipien von Privacy by Design nach [9]

Prinzip	Erläuterung
1. **Proaktiv** statt reaktiv; **Prävention** statt Schadensbegrenzung	Der PbD-Ansatz zeichnet sich durch proaktive und nicht durch reaktive Maßnahmen aus. Er antizipiert und verhindert in die Privatsphäre eingreifende Ereignisse, bevor sie eintreten. PbD wartet nicht darauf, dass Risiken für die Privatsphäre eintreten, und bietet auch keine Lösungen für Datenschutzverletzungen an, wenn diese bereits eingetreten sind - es zielt darauf ab, sie zu verhindern. Kurz gesagt: PbD kommt vor der Tat, nicht danach.
2. Datenschutz als **Standardeinstellung**	Mit PbD soll ein Höchstmaß an Privatsphäre gewährleistet werden, indem sichergestellt wird, dass personenbezogene Daten in jedem IT-System und jedem Geschäftsprozess automatisch geschützt werden. Wenn der Einzelne nichts unternimmt, bleibt seine Privatsphäre trotzdem gewahrt. Der Einzelne muss nichts unternehmen, um seine Privatsphäre zu schützen - sie ist standardmäßig in das System eingebaut.
3. Datenschutz **ist eingebettet** ins Design	PbD ist in den Entwurf und die Architektur von IT-Systemen und Geschäftspraktiken eingebettet. Er wird nicht im Nachhinein als Zusatz hinzugefügt. Das Ergebnis ist, dass der Datenschutz ein wesentlicher Bestandteil der bereitgestellten Kernfunktionalität wird. Der Datenschutz ist ein integraler Bestandteil des Systems, ohne die Funktionalität zu beeinträchtigen.
4. Volle Funktionalität – **Kein Nullsummenspiel**	PbD versucht, alle legitimen Interessen und Ziele auf eine positive „win-win"-Weise in Einklang zu bringen, nicht durch einen veralteten Nullsummen-Ansatz, bei dem unnötige Kompromisse eingegangen werden. PbD vermeidet die Vorspiegelung falscher Dichotomien, wie z.B. Privatsphäre vs. Sicherheit, und zeigt, dass es möglich und weitaus wünschenswerter ist, beides zu haben.
5. Von Anfange bis Ende Sicherheit – **Voller Life-Cycle-Schutz**	PbD, wurde bereits vor der ersten Datenerfassung in das System integriert und erstreckt sich somit über den gesamten Lebenszyklus der betroffenen Daten - starke Sicherheitsmaßnahmen sind für den Datenschutz von Anfang bis Ende unerlässlich. Dadurch wird sichergestellt, dass alle Daten sicher aufbewahrt und am Ende des Prozesses sicher und rechtzeitig vernichtet werden. Auf diese Weise gewährleistet PbD ein sicheres Life-Cylce-Management von Informationen von der Wiege bis zur Bahre, von Anfang bis Ende.
6. **Visibilität** und **Transparenz** – **Offenheit** wahren	PbD soll allen Beteiligten die Gewissheit geben, dass unabhängig von der jeweiligen Geschäftspraxis oder Technologie die angegebenen Versprechen und Ziele tatsächlich eingehalten werden, was von unabhängiger Seite überprüft wird. Die einzelnen Bestandteile und Abläufe bleiben für Nutzer und Anbieter gleichermaßen sichtbar und transparent. Denken Sie daran: Vertrauen Sie, aber überprüfen Sie!
7. **Respektieren** der Privatsphäre der Nutzer – **Nutzerzentriert** bleiben	In erster Linie verlangt PbD von Architekten und Betreibern, dass sie die Interessen des Einzelnen in den Vordergrund stellen, indem sie Maßnahmen wie starke Datenschutzvorgaben, angemessene Hinweise und benutzerfreundliche Optionen anbieten. Der Nutzer muss im Mittelpunkt stehen!

Situationen geeigneter erscheinen, wie zum Beispiel [342, 343]. Wiederum andere Studien untersuchen, das Verhalten von Nutzern in Bezug auf Ihre Privatsphäre [109, 113] oder befragen Personen hinsichtlich Ihrer Bedenken und Sorgen was ihre Privatsphäre angeht [344, 345, 346, 347, 348]. Nach meinem besten Wissen, existierten aber keine Untersuchungen darüber, wie Nutzer auf gezielte Erklärungen zu Datenpraktiken reagieren (sogenannte *Privacy Explanations*, siehe Abschnitt 6.1.3) und welchen Einfluss diese auf das Vertrauensverhältnis zwischen Endbenutzer und Informationssystem haben.

Bevor ich nachfolgend auf das angewandte Forschungs-Design und die Forschungsergebnisse zu sprechen komme, möchte ich im Vorfeld nochmal die Begriffe *Privatsphäre* und *Online-Privatsphäre* aufgreifen, voneinander differenzieren, definieren und somit für den SE-Kontext greifbarer und anwendbarer machen.

Zugehörige Publikationen Der in diesem Kapitel vorgestellte Forschungsbeitrag entstand in Kollaboration[3] mit vier anderen Forschern: Larissa Chazette, Kai Korte, Kurt Schneider und Alexander Specht. Dieses Kapitel gründet auf der gemeinsamen Zusammenarbeit und die Ergebnisse dazu wurden in [41] und [118] veröffentlicht, worauf dieses Kapitel basiert.

6.1 Privatsphäre und Erklärbarkeit

Der folgende Abschnitt beleuchtet die Konzepte von Privatsphäre und Erklärbarkeit. Dabei grenze ich den Begriff der Online-Privatsphäre und der Bedeutung für das SE ab und schließe mit der Verbindung beider Konzepte.

6.1.1 Das Konzept der Privatsphäre

Obwohl Forscher verschiedenster Disziplinen das Konzept der Privatsphäre aus unterschiedlichen Blickwinkeln untersucht haben, gibt es immer noch keine einheitliche Sichtweise oder Definition dieses Konzepts [349, 350, 351, 352, 353, 354].

Ein Überblick über die verschiedenen Interpretationen und Gedanken zum Konzept und Begriff der „*Privacy*“ sowie ihrem Einfluss und Auswirkungen auf Individuen gebe ich in Tabelle 6.2.[4] Was all diese Meinungen und Aussagen gemeinsam haben ist, dass es bei der Privatsphäre um die Kontrolle von Informationen über persönliche Angelegenheiten und die Schaffung privater Räume geht, sei es physischer oder mentaler Art. Ein physischer privater

[3] Aufgrund der Kollaboration mit meinen Forschungskollegen verwende ich das „wir“.

[4] Ich habe die Zitate bewusst nicht ins Deutsche übersetzt, um die von den Autoren erdachte Bedeutung nicht zu verzerren.

Tabelle 6.2: Auszug über die Interpretationen einiger Autoren zum Thema Privatsphäre und ihre Auswirkungen auf Individuen

Interpretationen der Autoren	Reference
Privacy is „a legal right“ and „the right to be let alone“	[355]
„Privacy is structured by the the answer […] to the questions ‚who are the persons you wish to exclude from having this knowledge?‘“	[356]
„Privacy is the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others“	[104]
Privacy is the right to select what personal information about an individual is known to what people	[104, 357]
„The desire by each of us for physical space where we can be free of interruption, intrusion, embarrassment, or accountability and the attempt to control the time and manner of disclosures of personal information about ourselves“	[358]
„The right to privacy exists because democracy must impose limits on the extent of control and direction that the state exercises over the day-to-day conduct of individual lives“	[105]
„A state in which persons may find themselves“	[106]
Influence of a person’s well-being	[107, 359, 350, 360]
It is important for our mental and physical health	[361]
Privacy is important to grow personally and its autonomy leaves room in determining one’s own path in life	[350]
Privacy is important for the endurance of a strong society	[107]
It acts like some sort of glue that binds individuals together in healthy relationships	[362]
A loss of privacy is not only unsettling but also represents an insult to a person’s dignity, independence, and integrity	[363, 364]
The scope of privacy is wide-ranging-potentially extending over information, activities, decisions, thoughts, bodies, and communication.	[365, 366]
„privacy is a sweeping concept, encompassing (among other things) freedom of thought, control over one’s body, solitude in one’s home, control over personal information, freedom from surveillance, protection of one’s reputation, and protection from searches and interrogations.“	[367]

Raum besteht aus baulich abgetrenntem Elementen (Decken, Wänden, Böden, Fenster und Türen). Das kann ein einzelnes Zimmer oder auch ein ganzes Gebäude sein. Ein privater Raum mentaler Art oder auch *persönlicher Raum* bezeichnet einen psychischen Aspekt eines Individuums, der auf die innere (mentale) Abgrenzung dieses Individuums gegenüber seiner Umwelt abzielt [368].

Die in Tabelle 6.2 erwähnten Konzepte zum Thema Privatsphäre der verschiedenen Autoren und Wissenschaftler fasse ich daher in nachfolgender Definition zusammen. Ziel dieser Definition ist *(a)* eine prägnante „Arbeitsdefinition“ zu schaffen, die die unterschiedlichen Nuancen der Privatsphäre erfasst und *(b)* ein gemeinsames Verständnis dafür entwickelt, was

Privatsphäre eigentlich bedeutet, um das Konzept im SE einfacher anwendbar zu machen sowie dem Mangel an Wissen zu diesem Qualitätsmerkmal zu begegnen [328, 369, 370]. Daher definiere ich Privatsphäre wir folgt:

Definition 6.1.1: Privacy

Privatsphäre ist ein Recht, ein Anspruch und ein Zustand, in dem ein Individuum unabhängig seine Grenzen festlegt und bestimmt, welche persönlichen Angelegenheiten[a] es mit anderen Individuen oder der Gesellschaft teilen oder (ihnen/dieser) vorenthalten möchte. Diese Grenzen können durch physische oder psychologische Mittel geschaffen werden[b]. Während dieses freiwilligen und vorübergehenden Rückzugs ist der Einzelne innerhalb dieser Grenzen vor unerwünschtem Eindringen (physisch oder mental), Peinlichkeit, Verurteilung, Diskriminierung, Verantwortlichkeit und gesellschaftlichen Normen und Zwängen geschützt.

[a] Persönliche Angelegenheiten beziehen sich auf persönliche Informationen, Gedanken, Gefühle und Gewohnheiten

[b] Private Räume, mentaler oder physischer Art.

Ungeachtet der Tatsache, dass Privatsphäre als ein Recht jedes Einzelnen betrachtet werden sollte, müssen darüber hinaus auch regulatorische Entitäten existieren, die die strikte Einhaltung dieses Rechts mit juristischen Mitteln überwachen und gewährleisten.

6.1.2 Online Privatsphäre - Bedeutung für das Software Engineering

Unser Umgang mit persönlichen Informationen musste aufgrund der rasanten Entwicklung der Informations- und Kommunikationstechnologie ständig angepasst und überarbeitet werden. So vereinfachte beispielsweise die Drucktechnik die Vervielfältigung und Verbreitung privater Informationen. Dieser Fortschritt der Digitalisierung und die Nutzung des Internets sorgen nun dafür, dass die Grenzen zwischen persönlichem und öffentlichem Leben immer stärker verschwimmen und sich verschieben. Wir kommunizieren über E-Mail, Messenger und soziale Medien, suchen über Suchmaschinen nach Antworten auf private und häufig sensible Fragen. Infolgedessen hinterlassen wir eine beträchtliche Menge an digitalen Spuren über unsere Abläufe und Routinen, Meinungen und Einstellungen, so dass unsere Privatsphäre stark davon abhängt, wie diese Systeme gestaltet sind [42, 9]

In der Offline-Welt – der physischen Welt – haben Individuen zumeist die Kontrolle über ihre eigene Privatsphäre und sind durch das normative Konzept der Privatsphäre geschützt, das durch soziale Normen und gesetzliche Grundlagen gestützt wird. Wenn sich beispielsweise

zwei Menschen auf einem öffentlichen Platz unterhalten, können herumstehende Dritte höchstens Teile der Konversation mithören. Die zwei Menschen haben aber auch die Möglichkeit sich weiter von den Umherstehenden zu entfernen, um sicherzustellen, dass ihr Gespräch unter ihnen – *privat* – bleibt. Also niemand mithören kann. In der Online-Welt ist diese Strategie jedoch nicht einfach so anwendbar und unwirksam. Natürlich gibt es Technologien wie Kryptographie, die eine verschlüsselte, nicht von Dritten abhörbare Kommunikation ermöglichen. Allerdings ist dies an dieser Stelle nicht gemeint, denn in der digitalen Welt besteht kein solcher Rückzug und „außer Hörweite" gehen, wie in der physischen Welt, denn allein schon die Menge und Einfachheit der hier zu speichernden Kommunikationsdaten ist wesentlich größer als bei einem kurzen Plausch auf einem öffentlichen Platz. Acquisti et al. geben zu bedenken, dass „Aktivitäten, die früher privat waren oder nur mit wenigen geteilt wurden, hinterlassen heute Datenspuren, die unsere Interessen, Eigenschaften, Überzeugungen und Absichten offenlegen" [371].

Um die Kontrolle über die Privatsphäre in der Online-Welt zu erlangen, müssen die Nutzer selbst aktiv für ihre Privatsphäre sorgen (sich selbst schützen), was wiederum ein Mindestmaß an technischem Wissen erforderlich macht (beispielsweise der Einsatz von Kryptographie als Security-Maßnahme). Nutzer können sich in der Online-Welt nicht auf Rechtssysteme verlassen und können auch nicht erwarten, dass andere Nutzer ihre sozialen und kulturellen Normen einhalten [351]. Effektiver Selbstschutz bedeutet, *(a)* sich über den Schutz der eigenen Privatsphäre bewusst zu sein und *(b)* entsprechende Maßnahmen zu ergreifen. Zudem wird dafür, wie angesprochen, ein gewisses technisches Wissen darüber vorausgesetzt, ***was*** schützenswert ist und ***wie*** man es tut. Jedoch verfügt nicht jeder über dieses notwendige Wissen, um Probleme mit der Privatsphäre zu erkennen und der aktive Schutz der eigenen Privatsphäre wird von vielen zudem auch als zu beschwerlich und zu zeitaufwändig angesehen [109].

Laut der Definition der Privatsphäre (Abschnitt 6.1.1), impliziert diese die Fähigkeit einer Person, ihre körperliche oder geistige Anwesenheit zu kontrollieren und ihr Recht auf geistige und/oder körperliche Abgeschiedenheit. Im Gegensatz dazu geht es bei der Online-Privatsphäre um die Kontrolle einer Person über seine persönlichen Informationen (Privatsphäreaspekte, siehe Abschnitt 2.5.3) im virtuellen Raum und sein Recht, diese Informationen bei Bedarf zurückzuhalten. Für die Entwicklung von Systemen und Technologien, die die Privatsphäre seiner Nutzer respektieren ist es daher entscheidend, zu verstehen was Online-Privatsphäre eigentlich im SE-Kontext bedeutet. Zu diesem Zweck definiere ich Online-Privatsphäre wie folgt:

Definition 6.1.2: Online Privacy

Online-Privatsphäre ist ein Recht, ein Anspruch und ein Zustand, in dem ein Individuum (Datenbesitzer) A seine Grenzen setzt und bestimmt (entscheidet oder kontrolliert), welchen *Privatsphäreaspekt*[a] X es mit Datensammlern und Datennutzern S teilen möchte, wer darauf zugreifen darf und zu welchem Zeitpunkt dies geschieht.

[a] Ein Privatsphäreaspekt (siehe Abschnitt 2.5.3) bezieht sich auf persönliche Informationen, Gedanken und Gefühle. In Bezug auf die Online-Privatsphäre, aber insbesondere auf Daten oder Informationen über eine Person. Beispiele: Name, Adresse, Bankdaten, GPS-Standort usw.

Der Vollständigkeit halber sei hier erwähnt, das die Online Privacy eines Individuums immer dann betroffen ist, wenn dieses Individuum mit einem digitalen Informationssystem interagiert, welches (persönliche) Daten dieser Interaktion erhebt, speichert und verarbeitet. Dies kann ein Auto, ein Fahrkartenautomat oder auch das Surfen im Internet sein. Allerdings muss technisch gesehen, nicht notwendigerweise ein Zugang zum Internet bestehen. Somit wäre der Begriff der *digitalen Privatsphäre* möglicherweise ebenfalls treffend, jedoch habe ich mich bewusst dagegen entschieden. Denn der Begriff der Online Privacy ist in der Literatur bereits etabliert, aber nicht definiert. Darüber hinaus sind im Zuge des Internet of Things (IoT) ohnehin ein Großteil der (Smart) Devices mit dem Internet verbunden.

6.1.3 Privacy Explanations

Erklärbarkeit wird als geeignetes Mittel angesehen, um der mangelnden Transparenz eines Systems entgegenzuwirken [166, 372, 373]. Darüber hinaus wirkt sich Erklärbarkeit auf das Vertrauensverhältnis zwischen Benutzer und System aus und kann zu einer höheren Akzeptanz beim Endnutzer führen [4]. Zu diesem Zweck habe ich das Konzept der Erklärbarkeit genutzt, um Endbenutzer in Bezug auf ihre Privatsphäre zu informieren und zu unterrichten. Dazu habe ich das Konzept der *Privacy Explanation* geschaffen. Ich verwende im Rahmen meiner Dissertation durchgehend den englischen Begriff, da die deutschen Begriffe „Erklärung zum Datenschutz“ oder auch „Privatsphäreerklärung“ irreführend sowie semantisch überladen sind und das dahinter liegende Konzept nicht korrekt widerspiegeln. In dieser Arbeit ist mit einer Privacy Explanation **keine** Datenschutzrichtlinie, Datenschutzerklärung oder auch **kein** Datenschutzhinweis im üblichen Sinne gemeint.

Ich nutze unsere Definition aus Abschnitt 4.2, um eine *Privacy Explanation* wie folgt zu definieren:

Definition 6.1.3: Privacy Explanation

Eine **Privacy Explanation** ist ein 6-Tupel $PE_x := (I, S, A, C, P, X)$. Hierbei ist I ein Korpus von Informationen, die ein System S einem Adressaten A im Kontext C gibt, um den Zweck P für die Verwendung eines *Privatsphäreaspekts* X zu erklären.

Die Erklärung I soll dem Endnutzer (Adressat A) eine Erklärung liefern, d.h. einen Grund, warum die Privatsphäre bezogenen Informationen des Nutzers (Privatsphäreaspekt X) benötigt werden (ein Grund könnte hier z.B. sein, ***warum*** eine Smartphone-App den Standort des Benutzers benötigt). Es ist notwendig, dass eine Privacy Explanation diese Begründung (den Zweck P) liefert, warum X benötigt wird und zwar auf transparente und verständliche Art und Weise. Diese Erklärung kann in Form von Text, visuell, auditiv oder in einer beliebigen Kombination dieser Elemente gegeben werden. Der Kontext ist die Situation, in der eine Erklärung gegeben wird, und besteht „aus der Interaktion zwischen einer Person, einem System, einer Aufgabe und einer Umgebung“ [4].

6.2 Forschungsvorgehen

Ziel unserer Forschung war es, das Konzept der Privacy Explanations zu untersuchen. Genauer gesagt, sollte die Wahrnehmungen und Meinungen von Endbenutzern in Bezug auf Privacy Explanations untersucht werden. Das Forschungsziel wurde auch hier entsprechend des Goal Definition Templates [53, 54] formuliert:

> ***Goal Definition:*** Wir **analysieren** die Wahrnehmungen und Bedürfnisse von Endnutzern in Bezug auf Erklärungen zum Datenschutz in Softwaresystemen **zu dem Zweck** der Untersuchung, ob Erklärungen das Vertrauensniveau **aus der Sicht von** Endnutzern beeinflussen könnten **im Kontext** eines Online-Fragebogens.

Abbildung 6.1 gibt einen Überblick über das Vorgehen. Als Forschungsinstrument kam hierbei eine Online-Umfrage (Survey) zum Einsatz, dessen Daten anschließend qualitativ und quantitativ ausgewertet wurden.

6.2.1 Survey Design

Um die Qualität unserer Umfrage zu gewährleisten, haben wir uns an etablierte Richtlinien für die Erstellungen von Umfragen gehalten [316, 317, 374]. Um mögliche Verzerrungen der Antworten zu vermeiden, sind wir den Richtlinien gefolgt und haben verschiedene Maßnahmen ergriffen, wie z.B. die Vermeidung von Leitfragen, verwendeten eine präzise und einfache

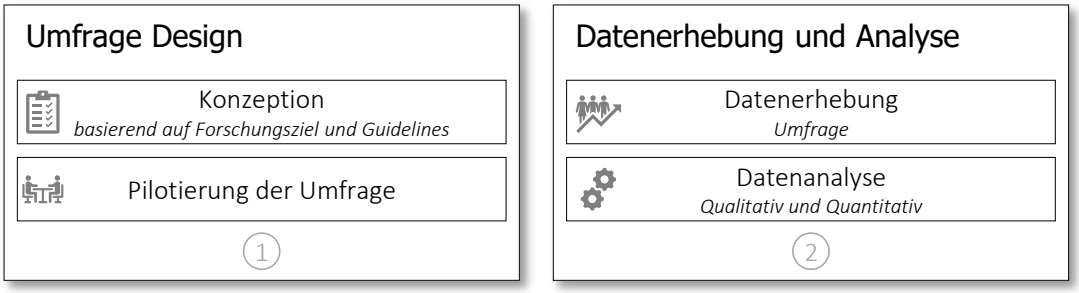


Abbildung 6.1: Überblick des Forschungsvorgehens

Sprache, hielten unsere Fragen kurz und klar und versuchten ausgewogene sowie einheitliche Antwortoptionen zu verwenden. Außerdem informierten wir die Teilnehmer in schriftlicher Form, dass die Umfrage anonym ist und sie ehrlich antworten sollten, da es keine richtigen oder falschen Antworten gibt. Die vollständige Umfrage ist in Abschnitt D.1 zu finden.

Wir haben die Struktur der Umfrage in Übereinstimmung mit unserem Forschungsziel konzipiert. Die Umfrage begann mit einer kurzen Einleitung und umfasste vier Blöcke, gegliedert in acht logische Teile mit insgesamt 34 Fragen (30 Multiple-Choice-Fragen, vier offene Fragen). Einige der Multiple-Choice-Fragen boten den Befragten auch die Möglichkeit, einen eigenen Antworttext zu formulieren, wenn ihnen keine der vorgegebenen Antwortmöglichkeiten zusagten. Die Struktur der Umfrage ist in Abbildung 6.2 abgebildet.

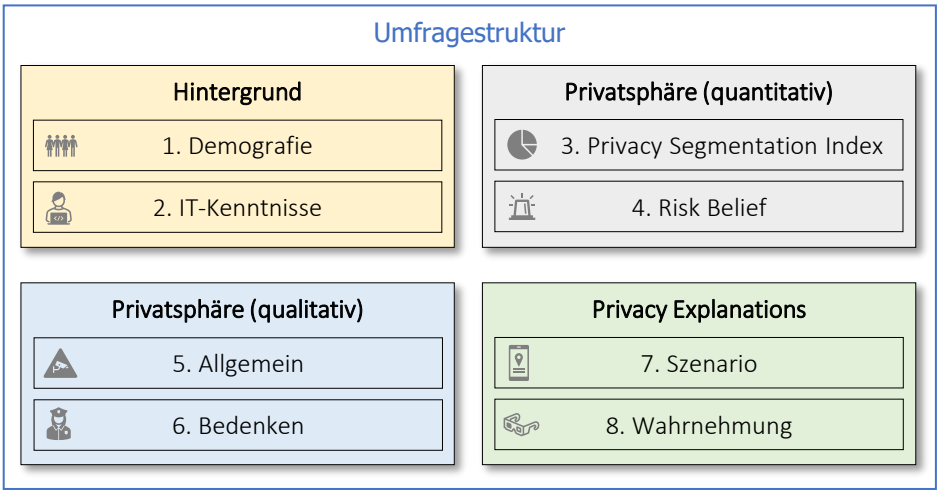


Abbildung 6.2: Überblick über die Struktur der Umfrage

In Teil 1 ging es darum, demografische Faktoren zu ermitteln, die einen möglichen Einfluss auf die Antworten der Befragten haben könnten. Der zweite Teil bestand aus Fragen, die das IT-Hintergrundwissen der Teilnehmer abfragte. Der dritte Teil enthielt Fragen zur Erfassung des PSI (Abschnitt 6.2.3.1) der Befragten. Im vierten Teil wurden Fragen zur Einschätzung des

erlebten Risikos (Risk Belief, Abschnitt 6.2.3.2) der Teilnehmer gestellt. Der dritte Block umfasste Fragen für eine qualitative Bewertung der Teilnehmer in Bezug auf ihre Privatsphäre. Hierzu wurden in Teil 5 allgemeine Fragen zum Privatsphäreverhalten der Teilnehmer gestellt und in Teil 6 fragten wir nach deren Bedenken. Die Teilnehmer wurden hier gebeten, Situationen zu beschreiben, in denen sie bei der Nutzung von Software Bedenken hinsichtlich ihrer Privatsphäre hatten. Der letzte Block befasste sich mit den Privacy Explanations. Hierfür gaben wir den Teilnehmern in Teil sieben folgendes hypothetisches Szenario:

> **Hypothetische Situation:** Sie haben über einem Bekannten von einer neuen App gehört, mit der Sie Sightseeing-Touren oder Tagesausflüge für Städte auf der ganzen Welt planen können. Sie beschließen, diese App für Ihren bevorstehenden Städte-Trip auf Ihr Smartphone herunterzuladen. Beim ersten Start der App fragt diese jedoch, ob Ihr Standort verwendet werden darf und fragt auch nach Ihrem Geburtsdatum. Sie sind sich über den Grund nicht sicher, da die App Ihnen keine weiteren Informationen über die Verwendung Ihrer Daten gibt.

Anschließend fragten wir die Teilnehmer zum einen „*Wie sehr sind Sie daran interessiert, eine Erklärung in Bezug auf Ihre Privatsphäre zu erhalten?*“, zum anderen präsentierten wir ihnen verschiedene Privacy Explanations und fragten sie nach ihrer Wahrnehmung. Im letzten Teil fragten wir die Teilnehmer ganz allgemein nach möglichen Vorteilen von Privacy Explanations und danach, wann ein System eine solche Erklärung geben sollte.

Um die Qualität der Umfrage sicherzustellen, haben wir vier Runden von Pilotierungen der Umfrage durchgeführt. Zwei davon fanden mit Mitgliedern unserer Forschungsgruppe und zwei mit Kandidaten der Zielpopulation statt. Auf der Grundlage dieser Pilotierungen haben wir einige kleinere Korrekturen vorgenommen wie z.B. Hinzufügen von Informationen zur Interpretation bestimmter Fragen oder kleinere Textänderungen zum besseren Verständnis.

6.2.2 Datenerhebung und Analyse

Die Daten wurden mittels eines web-basierten Fragebogens erhoben. Die Umfrage wurde mit dem Umfragetool LimeSurvey[5] erstellt und auf den Servern der Leibniz Universität gehostet.

6.2.2.1 Datenerhebung

Die Datenerhebung erfolgte über zwei Monate, beginnend im Juni 2021. Wir verbreiteten die Umfrage über verschiedene Wege, darunter akademische Mailing-Listen, Facebook und Twitter, und wir ermunterten unser persönliches Netzwerk ein, die Umfrage in ihren Netzwerken

[5] `https://www.limesurvey.org`

ebenfalls zu teilen. Unsere Zielgruppe waren volljährige Endnutzer aus unterschiedlichen Berufen und mit variierenden IT-Kenntnissen, da wir die Wahrnehmung von Endnutzern mit unterschiedlichem Hintergrund zu diesem Thema verstehen wollten. Auf Basis unserer Sampling-Strategie (Kontaktnetzwerke vor allem in Deutschland und Brasilien) stellten wir unsere Umfrage in drei verschiedenen Sprachen zur Verfügung: Englisch, Deutsch und Portugiesisch. Wir gingen davon aus, dass ein großer Teil der Teilnehmer aus Brasilien und Deutschland kommen würde.

6.2.2.2 Analyse

Für die Analyse und Auswertung der Umfrageergebnisse nutzten wir qualitative und quantitative Analysetechniken. Wir haben die Ergebnisse in Tabellenkalkulationen exportiert, um deskriptive Statistiken zu berechnen. Für die offenen Fragen wendeten wir eine qualitative Datenanalyse an, die aus einem offenen Kodierungsansatz bestand, wie von Saldaña [230] beschrieben. Laut Saldaña ist das Kodieren „eine Art der Analyse qualitativer Daten". Es „transformiert qualitative Daten in quantitative Daten, aber es beeinträchtigt nicht ihre Subjektivität oder Objektivität" [375]. Unser Kodierungsprozess folgte ebenfalls wie in Abschnitt 4.1.2.2 und Absatz 5.1.3.3 zwei konsekutiven Kodierungszyklen, bestehend aus In Vivo Coding [231] und Pattern Coding [232].

Der Kodierungsprozess wurde von meinem Kollegen Alexander Specht und mir unabhängig voneinander durchgeführt. Im Falle von Diskrepanzen diskutierten wir die Unterschiede, bis wir einen Konsens erreichten. Zur Bewertung der Zuverlässigkeit des Kodierungsverfahrens haben wir die Cohens-Kappa-Statistik [376] verwendet. Der daraus resultierende Wert von $\kappa = 0,87$ zeigte eine nahezu perfekte Übereinstimmung [228].

6.2.3 Privacy Segmentation Index und Risk Belief

In unserer Umfrage haben wir die Teilnehmer auf der Grundlage des Privacy Segmentation Index (PSI) klassifiziert und haben ihr erlebtes Risiko (Risk Belief) in Bezug auf ihre Privatsphäre ermittelt. Im Folgenden beschreibe ich was PSI und Risk Belief sind und wie wir diese Metriken erhoben haben.

6.2.3.1 Privacy Segmentation Index

Westin entwickelte den PSI, um die Verbraucher nach ihren Bedenken hinsichtlich des Schutzes der Privatsphäre zu klassifizieren [344]. Obwohl sich der PSI auf die Verbraucherperspektive bezieht, wird er in einem breiteren Kontext angewandt [33, 377, 378]. Vor diesem

Hintergrund haben wir ebenfalls den PSI erhoben, um die Haltung der Teilnehmer hinsichtlich ihrer Datenschutzbedenken zu erfassen. Der PSI beinhaltet folgende Aussagen, wobei die Befragten ihre Zustimmung auf einer 7-stufigen Likert-Skala (*stimmt überhaupt nicht zu* bis *stimme völlig zu*) zum Ausdruck bringen:

P1. Die Verbraucher haben jegliche Kontrolle darüber verloren, wie persönliche Informationen von Unternehmen gesammelt und verwendet werden.

P2. Die meisten Unternehmen gehen mit den persönlichen Daten, die sie über Verbraucher sammeln, ordnungsgemäß und vertraulich um.

P3. Bestehende Gesetze und Organisationspraktiken bieten heute ein angemessenes Maß an Schutz für die Privatsphäre der Verbraucher.

Anhand der Antworten können die Teilnehmer dann einer von drei Kategorien zugeordnet werden: *Privacy Fundamentalists*, *Privacy Pragmatists* oder *Privacy Unconcerned*. Die Beschreibungen dieser Kategorien werden im Harris-Bericht von Westin wie folgt wiedergegeben [379]:

> **Privacy Fundamentalists:** Diese Gruppe misst die Privatsphäre einen besonders hohen Wert bei, lehnt die Behauptung vieler Organisationen ab, dass sie persönliche Informationen für ihre Geschäfts- oder Regierungsprogramme benötigen oder dazu berechtigt sind, und ist der Meinung, dass mehr Einzelpersonen die Herausgabe von Informationen, nach denen sie gefragt werden, einfach verweigern sollten, und befürwortet den Erlass strenger Bundes- und Landesgesetze zur Sicherung der Datenschutzrechte und zur Kontrolle der Ermessensfreiheit von Organisationen.

> **Privacy Pragmatists:** Diese Gruppe wägt den Wert verschiedener Geschäfts- oder Regierungsprogramme, die persönliche Informationen erfordern, für sie und die Gesellschaft ab, prüft die Relevanz und die gesellschaftliche Akzeptanz der gesuchten Informationen, möchte die potenziellen Risiken für die Privatsphäre oder die Sicherheit ihrer Informationen kennen, prüft, ob faire Informationspraktiken auch ausreichend beachtet werden und entscheidet dann, ob sie bestimmten Informationsaktivitäten zustimmt oder sie ablehnt — wobei ihr Vertrauen in die betreffende Branche oder das betreffende Unternehmen ein wesentlicher Entscheidungsfaktor ist. Die Pragmatiker bevorzugen freiwillige Standards und die Wahlfreiheit der Verbraucher gegenüber Gesetzen und staatlicher Durchsetzung. Aber sie werden die Gesetzgebung unterstützen, wenn sie der Meinung sind, dass auf freiwilliger Basis nicht genug - oder nichts sinnvoll erscheinendes - getan wird.

> **Privacy Unconcerned:** Diese Gruppe versteht nicht, was es mit der Aufregung um den Datenschutz auf sich hat, befürwortet die Vorteile der meisten organisatorischen Programme mehr als die Warnungen vor dem Missbrauch der Privatsphäre, hat wenig Probleme damit, ihre persönlichen Daten an Regierungsbehörden oder Unternehmen weiterzugeben, und sieht keine Notwendigkeit, eine weitere Regierungsbürokratie (einen „Big Brother" auf Bundesebene) zu schaffen, um die Privatsphäre von Personen zu schützen.

Laut Westin erfolgt die Klassifizierung wie folgt: *Privacy Fundamentalists* stimmen mit P1 überein und stimmen sowohl P2 als auch P3 nicht zu. Teilnehmer, die mit P1 nicht einverstanden sind und mit P2 und P3 einverstanden sind, gehören zur Klasse der *Privacy Unconcerned*. Die übrigen Teilnehmer können als *Privacy Pragmatists* eingruppiert werden.

6.2.3.2 Risk Belief

Der Risk Belief als Metrik lässt sich auf Tsai et al. [345] zurückführen. Mit Hilfe dieser Metrik lässt sich das empfundene Risiko quantifizieren, das eine Person durch die Weitergabe ihrer Informationen im Internet wahrnimmt. Wir haben den Risk Belief von Tsai et al. für unsere Umfrage übernommen, um für jeden Teilnehmer eine Risikobewertung zu berechnen. Unser Ziel war es, den Risk Belief zu nutzen, um eine bessere Einschätzung der einzelnen Teilnehmer zu erhalten. Zu diesem Zweck haben wir die Fragen zum Risk Belief etwas modifiziert und die Fragen nicht nur in Bezug auf das Online-Shopping, sondern in Bezug auf Online-Dienste im Allgemeinen formuliert. Zu diesem Zweck haben wir den Befragten insgesamt vier geschlossene Fragen gestellt:

Q1. Ich fühle mich sicher, bei der Weitergabe meiner persönlichen Daten an Online-Dienste (wie Online-Shops) und/oder Apps.*

Q2. Die Bereitstellung von persönlichen Informationen an Online-Dienste oder Apps verursacht zu viele Bedenken.

Q3. Im Allgemeinen vertraue ich Online-Unternehmen im Umgang mit meinen persönlichen Daten, z.B. meiner Shopping-Historie.*

Q4. Wie besorgt sind Sie über die Bedrohungen Ihrer persönlichen Privatsphäre im Internet heutzutage?

Für die Fragen eins bis drei verwendeten wir eine 7-Punkte Likert-Skala, um den Grad der Zustimmung zu ermitteln (*stimme überhaupt nicht zu* bis *stimme völlig zu*). Bei der vierten Frage wurde der Grad der Besorgnis mit einer 5-Punkte Likert-Skala (*überhaupt nicht besorgt* bis *extrem besorgt*) gemessen. Die Antworten wurden entsprechend der Skala bewertet,

wobei die Werte für die Fragen eins und drei umgekehrt wurden, um das Gefühl der Besorgnis widerzuspiegeln (je höher der Wert, desto höher das wahrgenommene Risiko). Um die 5-Punkte-Likert-Skala auf die 7-Punkte-Likert-Skala abzubilden, haben wir die Elemente mit 1,5 gewichtet. Die Verwendung einer 7-Punkte-Likert-Skala hier resultiert dadurch, dass möglichst konsistente Skalen innerhalb eines Surveys verwenden werden sollten. Wir halten diese Zuordnung für gerechtfertigt und die Aussagen der Teilnehmer für nicht verzerrt. Ein Chronbach's α-Wert [380] von $0,76$ bestätigte die Zuverlässigkeit der 7-Punkte-Skala [381].

6.3 Ergebnisse der Umfrage

In dieser Sektion präsentiere ich die analysierten Daten und Ergebnisse unserer Umfrage. Die Gliederung orientiert sich hierbei an der Struktur der Umfrage. Hierfür gebe ich zunächst in Abschnitt 6.3.1 die demografischen Daten wider. Anschließend gehe ich auf die Bedenken und Sorgen der Teilnehmer in Bezug auf Ihre Privatsphäre in Abschnitt 6.3.2 ein und lege dann die Ergebnisse hinsichtlich der Privacy Explanations dar (Abschnitt 6.3.3). Eine Interpretation und Diskussion der Ergebnisse, was abschließend auch zur Beantwortung von RQ2 führt, erfolgt in Abschnitt 6.5.

Insgesamt haben 209 Personen an unserer Umfrage teilgenommen. 155 von ihnen haben die Umfrage vollständig abgeschlossen. Für unsere Datenanalyse haben wir ausschließlich die 155 vollständigen Antworten berücksichtigt.

6.3.1 Demografie

Der Großteil der Teilnehmer stammte aus Deutschland (67,1%) und Brasilien (21,9%). Das Alter reichte von 19 bis 92 Jahren (M=39, SD=14,1). 61,9% der Teilnehmer bezeichneten sich selbst als männlich und 37,5% als weiblich. Eine Person (0,6%) ordnete sich anderweitig zu. Die Mehrheit der Teilnehmer arbeitete in den Bereichen Naturwissenschaften, Geografie und Informatik (32,03%), wie in Abbildung 6.3 dargestellt. Zwei der Befragten äußerten sich nicht zu der Fragen nach Ihrem Berufsfeld.

IT-Kenntnisse Um einen Eindruck über den Kenntnisstand der Teilnehmer zu erhalten in Bezug auf deren IT-Wissen, haben wir verschiedene Fragen zur Selbsteinschätzung in unsere Umfrage aufgenommen. Hier sollten die Befragten angeben, ob sie in der Lage sind, bestimmte Aufgaben auszuführen und ob sie mit bestimmten IT-Begriffen vertraut sind. Zur Eingruppierung unserer Teilnehmer haben wir vier Cluster gebildet aus *Experten*, *Entwicklern*, *Fortgeschrittenen* und *Anfängern*. *Experten* zeichnen sich durch hohes IT-Hintergrundwissen,

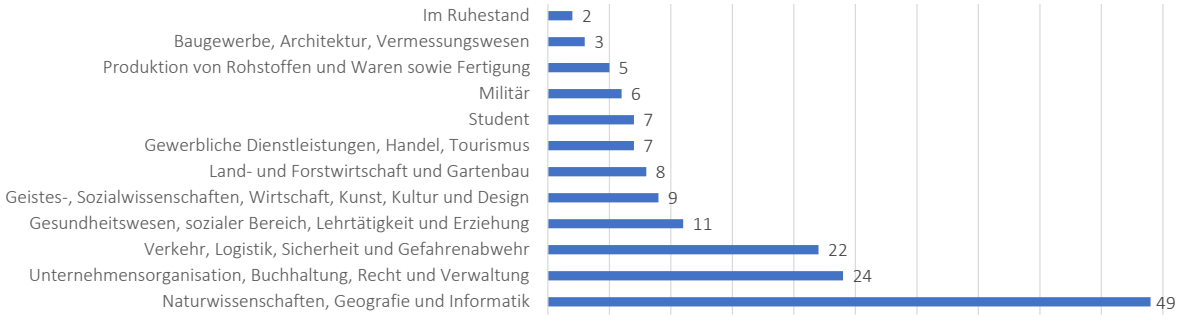
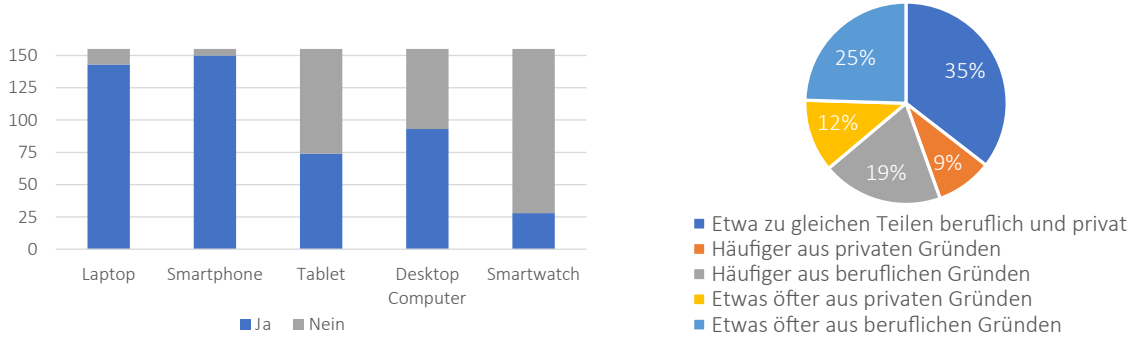


Abbildung 6.3: Überblick über die Tätigkeitsfelder der Teilnehmer

Kenntnisse im Umgang mit Software etc. aus, verfügen aber nicht notwendigerweise über Programmierkenntnisse. Das heißt, *Entwickler* verfügen also neben ihrer Programmierkenntnisse auch noch über profunde Kenntnisse über interne Systemaspekte. Wir haben die Befragten in die Gruppe *Fortgeschritten* eingeteilt wenn sie angaben, mindestens über gute Software-Kenntnisse zu verfügen (z.B. funktionale Tabellenkalkulationen zu erstellen oder in der Lage zu sind, neue Programme schnell zu erlernen) sowie über Kenntnisse grundlegender Computer-Konzepte. Befragte, die nur die Aussage „Ich habe nicht so viel Erfahrung, aber ich kann meine E-Mails abrufen und einfache Aufgaben mit Textverarbeitungsprogrammen erledigen“ als Antwort wählten, ordneten wir der Gruppe *Anfänger* zu. Insgesamt ordneten wir 66,5% der Teilnehmer den Clustern *Experte* und *Entwickler* zu (40 Teilnehmer Entwickler, 63 Teilnehmer Experten). 33 Teilnehmer (21,3%) wurden als *Fortgeschritten* und 19 (12,2%) als *Anfänger* eingestuft.

Nutzung von Geräten und Software Auf die Frage nach der Nutzung, welche digitalen Geräte die Teilnehmer verwenden, gaben 96,8% der Befragten an, dass sie im Alltag ein Smartphone verwenden, und alle Befragten nutzen täglich entweder einen Laptop oder einen Desktop-Computer (siehe Abbildung 6.4a).



(a) Nutzung digitaler Geräte

(b) Nutzung von Software

Abbildung 6.4: Nutzung von digitalen Geräten und Software

Auf die Frage, ob sie Softwaresysteme eher aus beruflichen oder privaten Gründen nutzen, gaben 35,5% der Befragten an, dass sie sie mehr oder weniger gleichermaßen aus beruflichen und privaten Gründen nutzen. 43,9% nutzen Software mehr für die Arbeit (19,4% häufiger für die Arbeit, 24,5% etwas häufiger für die Arbeit), wie in Abbildung 6.4b dargestellt. Beide Faktoren (Nutzung verschiedener Geräte und Nutzung von Software) zeigen, dass Softwaresysteme einen festen Platz im Alltag der Befragten einnehmen und deuten auf eine für unsere Zwecke aussagekräftige Population hin.

6.3.2 Sorgen in Bezug auf die Privatsphäre

Um die Bedenken und Sorgen der Befragten in Bezug auf ihre Privatsphäre zu ermitteln, klassifizierten wir die Teilnehmer gemäß dem PSI, wie in Abschnitt 6.2.3.1 beschrieben. Anschließend haben wir den Grad des Risikos quantifiziert, den die Befragten bei der Weitergabe ihrer Informationen im Internet wahrnehmen (siehe Risk Belief, Abschnitt 6.3.2.3). Abschließend wollten wir zudem wissen, welche Bedenken die Teilnehmer hinsichtlich ihrer Privatsphäre haben und welchen Risiken bzw. Bedrohungen sie sich ausgesetzt sehen.

6.3.2.1 Privacy Segmentation Index und Privatsphäreverhalten

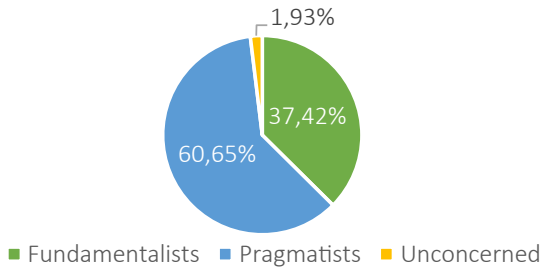


Abbildung 6.5: PSI-Verteilung der Teilnehmer

Abbildung 6.5 zeigt die PSI-Verteilung der Teilnehmer. Die Verteilung stimmt weitgehend mit den Daten aus Westins Umfragen zum Datenschutz (1996, 2000, 2001 und 2003) [344, 382] überein, bei denen die Mehrheit als Privacy Pragmatists eingestuft wurde. Die Anteile der Fundamentalists und Unconcerned unterscheiden sich in unseren Umfrageergebnissen etwas. Sie zeigen einen höheren Anteil von Fundamentalists und einen geringen Anteil von Unconcerned. Ob aus den Privacy Unconcerned nun Pragmatists oder Fundamentalists geworden sind, lässt sich auf der Grundlage unserer Daten nicht sagen. Der höhere Anteil der Fundamentalists könnte möglicherweise dadurch erklärt werden, dass die Menschen ihre (Online-) Privatsphäre möglicherweise mehr gefährdet sehen, als es in der Vergangenheit der Fall war. Auch die geografische Verteilung kann hier eine Rolle spielen, denn die Deutschen sind durchaus *privacy aware* und um ihre Online-Privatsphäre besorgt [383]. Aber auch hier ist es nicht möglich, diese Annahme allein auf Basis unserer Daten zu belegen.

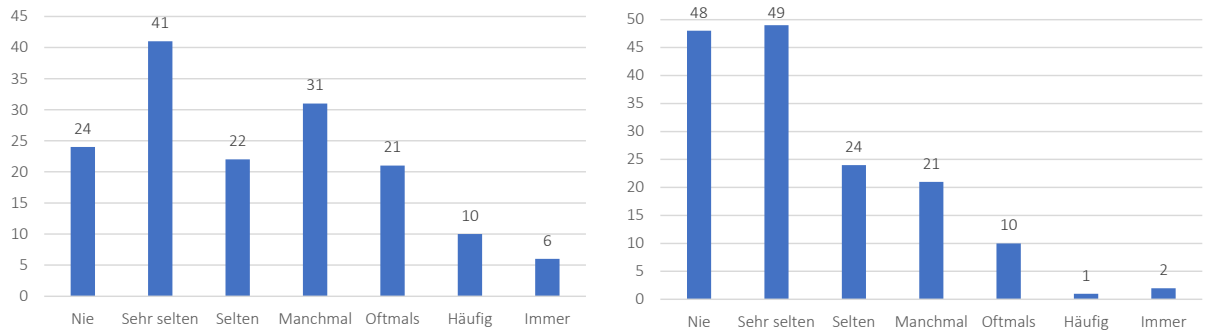
Erforderliche Berechtigungen bei Installation von Software Im Anschluss an die Fragen zum PSI haben wir die Teilnehmer gefragt, ob sie bei der Installation von Anwendungen auf „erforderliche Berechtigungen“ (engl.: *required premissions*) achten. 60% gaben an, dass sie immer darauf achten, 34,2% tun dies manchmal, und 5,8% achten nicht darauf. Auf die Frage, wie schnell sie bei der ersten Verwendung von Software die Schaltfläche „Zustimmen“ für die Geschäftsbedingungen drücken, gaben 30,3% der Befragten an, dass sie die Schaltfläche instantan drücken, 55,5% innerhalb einer Minute und 14,2% brauchen mehr als eine Minute, bevor sie die Schaltfläche drücken.

50,0% der Fundamentalists und 65,9% der Pragmatists achten immer auf die erforderlichen Berechtigungen. Allerdings gaben zwei als Unconcerned klassifizierte Personen (66,6%) an, dass sie auch immer darauf achten (der hohe Prozentsatz ergibt sich aus der Tatsache, dass von den 155 Befragten insgesamt 3 als Unconcerned eingestuft wurden). Der Anteil derjenigen, die gelegentlich darauf achten, ist bei den Fundamentalists (41,3%) höher als bei den Pragmatists (29,8%).

Auf die Frage, wie schnell sie bei der Installation von Software auf „Zustimmen“ drücken, ist der Prozentsatz, der dies *innerhalb einer Minute*$^{\diamond}$ oder *mehr als eine Minute** tut, bei den Pragmatikern ebenfalls etwas höher (57. 5%$^{\diamond}$, 14.9%*) als unter Fundamentalisten (53.5%$^{\diamond}$, 13.8%*), Privatsphäre Unbekümmert (33.3%$^{\diamond}$, 0%*).

Obwohl es keinen signifikanten Unterschied zwischen den Gruppen gibt, deuten die Ergebnisse darauf hin, dass sich die Nutzer der Risiken für die Privatsphäre in Bezug auf die erforderlichen Berechtigungen von Apps bewusst sind, da die Mehrheit darauf achtet, welche Berechtigungen eine App benötigt. Das spiegelt sich also in ihrem Verhalten wider. Auch bei der Installation von Software stimmt die Mehrheit nicht sofort zu. In unserer Umfrage wird nicht gefragt, ob die Nutzer in dieser Zeit zum Beispiel Informationen über den Datenschutz lesen. Wenn diese Daten jedoch in Verbindung mit der Frage nach den erforderlichen Berechtigungen analysiert werden, können wir daraus schließen, dass sich die Befragten der Gefahren für die Privatsphäre, die von Software ausgehen können, durchaus bewusst sind.

Datenschutzrichtlinien Damit wir in Bezug auf das Privatsphäreverhalten und der Haltung der Teilnehmer ein umfassenderes Bild zeichnen können, fragten wir, ob die Teilnehmer im Allgemeinen darauf achten, ob ein Anbieter einer digitalen Dienstleistung eine Datenschutzerklärung (DSE) zu Verfügung stellt und ob die Teilnehmer diese ggf. auch lesen. Die Daten dazu sind in Abbildung 6.6 dargestellt.
Die Ergebnisse decken sich mit denen aus der Umfrage aus Kapitel 5 sowie bereits erwähnter Literatur [42, 46, 48]. Die Gründe dafür sind ebenfalls bereits diskutiert worden. Dennoch war die Einbindung dieser Frage durchaus berechtigt, da es zum einen ein gutes Bild zum Verhalten und Umgang unserer Teilnehmer gibt hinsichtlich dem Schutz ihrer privaten Daten und zum



(a) Achten auf DSEs

(b) Lesen von DSEs

Abbildung 6.6: Beachtung von Datenschutzerklärungen (DSEs)

anderen ist die Sample-Größe dieser Umfrage fast dreimal so groß wie die aus Kapitel 5 und unterstreicht somit umso mehr die Dringlichkeit, alternative Lösungen zur Informierung und Aufklärung von Endbenutzern in Hinblick auf Datenpraktiken zu erforschen.

6.3.2.2 Risk Belief

Die berechneten Risk Scores für die Risk Belief-Metrik reichten von 1,95 bis 7,0 (M=4,78, SD=1,19). Das in Abbildung 6.7 dargestellte Histogramm zeigt eine annähernd normale Verteilung für die Risk Scores. Eine Prüfung auf Normalverteilung mit Shapiro-Wilk-Test [384] bestätigte, dass die Risk Scores normalverteilt sind (W=0,98, p=0,058). Die Mehrheit der Befragten (68,4%) hat einen Risikowert $> 4,3$. Dies deutet darauf hin, dass sich die Befragten nicht nur bewusst sind, dass sie ihre Daten online weitergeben, sondern dass sie dies auch als ein ziemlich hohes Risiko ansehen.

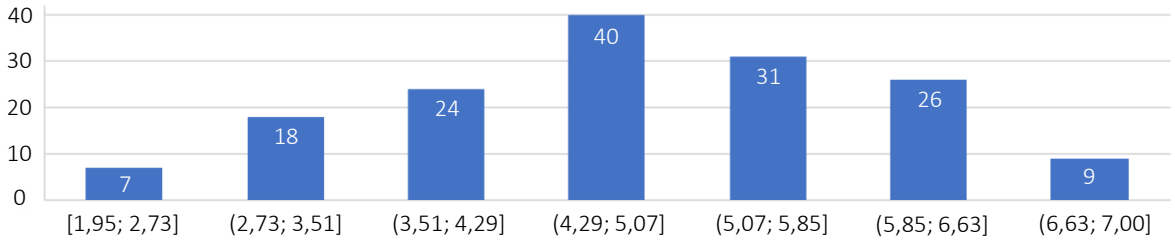


Abbildung 6.7: Histogramm der Risk Scores

Wir konnten beobachten, dass die Befragten mit einem höheren Risk Score eher DSEs lesen, was mit einer positiven Korrelation durch die Rangkorrelation nach Spearman [385] nachgewiesen werden konnte (ρ=0,41, p<0.001). Analog dazu fühlten sich Teilnehmer mit einem höheren Risk Score auch unwohler bei bei der Nutzung von Shopping-Portalen (ρ=0.52, p<0.001), Benutzung von Suchmaschinen (ρ=0.51, p<0.001) und Spielen (ρ=0.45, p<0.001) unwohl fühlten (vgl. Abbildung 6.8).

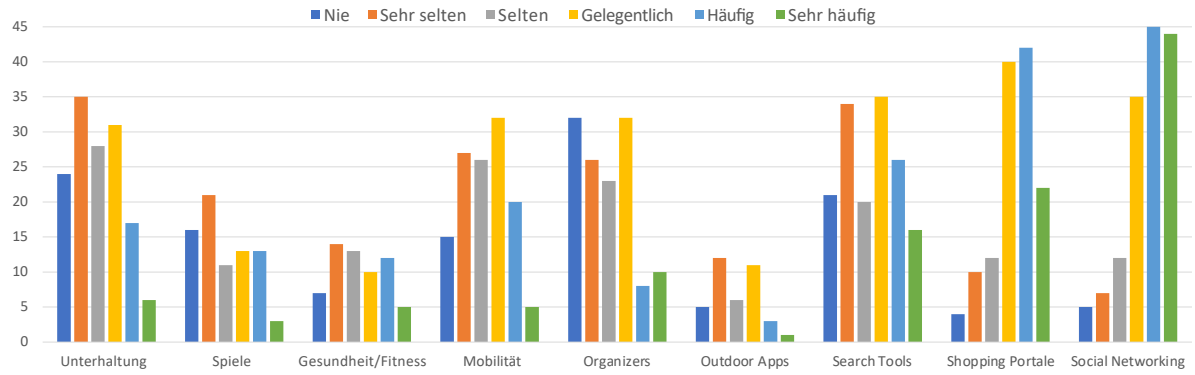


Abbildung 6.8: Wie oft fühlen Sie sich wegen des Datenschutzes unwohl, wenn Sie Software verwenden oder Websites besuchen, die mit diesen Kategorien in Zusammenhang stehen?

Wir analysierten auch den möglichen Zusammenhang zwischen dem IT-Kenntnisstand und dem PSI. Nach dem PSI ist der Risk Score bei den Privacy Fundamentalists im Durchschnitt etwas höher (M=5,52, SD=0,91) als bei den Pragmatists (M=4,36, SD=1,07). Der Risk Score Unconcerned (drei Befragte) ist im Durchschnitt am niedrigsten (M=3,63, SD=1,85). Dieser Unterschied zwischen den Gruppen ist statistisch signifikant laut Rangkorrelation nach Spearman mit einer negativen Korrelation (ρ=-0.49, p<0.001).

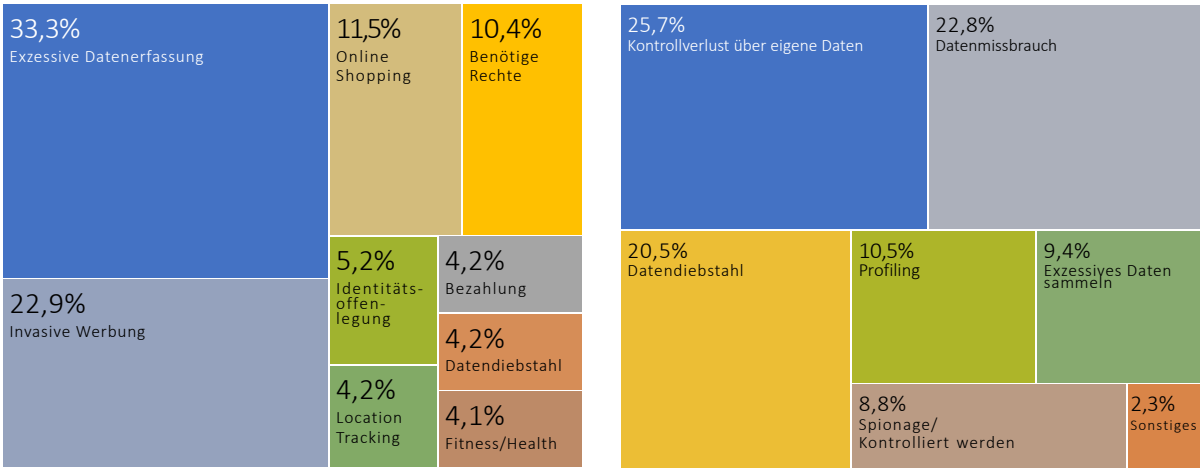
Innerhalb der IT-Gruppierungen gibt es keine statistisch signifikanten Unterschiede. Unseren Ergebnissen zufolge haben Fortgeschrittene (M=5,01, SD=1,11) und Anfänger (M=5,13, SD=0,94) jedoch einen durchschnittlich höheren Risk Score als Experten (M=4,55, SD=1,14) und Entwickler (M=4,77, SD=1,34). Diese Unterschiede sind aber geringer als zwischen den PSI-Gruppen. Möglicherweise ist der niedrigere Risk Score bei den Experten und Entwicklern darauf zurückzuführen, dass sie ein tieferes Verständnis von Softwaresystemen haben (insbesondere die Entwickler). Sie haben also eine Vorstellung davon, was Software „im Inneren“ tut. Fortgeschrittene und Anfänger hingegen sehen Softwaresysteme in erster Linie als Blackboxen sie haben also keine bis wenig Kenntnisse über möglicher interne Abläufe. Dieser Mangel an Wissen könnte der Grund für das höhere wahrgenommene Risiko sein. Diese Annahme lässt sich jedoch mit unseren Daten nicht begründen, da wir keine weiteren Fragen in diese Richtung gestellt haben.

6.3.2.3 Wahrgenommene Bedrohungen und Bedenken hinsichtlich der eigenen Privatsphäre

Damit Software Engineers auch für eine nutzerzentrierte Lösung zum Schutz und zur Aufklärung der Endbenutzer sorgen und beitragen können, ist es wichtig zu erfahren, was genau die wahrgenommenen Bedrohungen der Endbenutzer gegenüber ihrer Privatsphäre sind und was in ihnen Unwohlsein bei der Nutzung von Software auslöst. Dazu baten wir die Befragten (a)

eine Situation zu nennen, in der sie sich bei der Verwendung von Software im Hinblick auf die Privatsphäre besonders unwohl gefühlt haben. Zudem baten wir sie (b) uns eine Bedrohung in Bezug auf ihre Privatsphäre zu nennen, die ihnen besonders viel Sorgen bereitet.

Unbehagen bei der Verwendung von Software Wir wollten herausfinden was ein Gefühl des Unwohlseins bei unseren Teilnehmern in Bezug auf ihre Privatsphäre auslöst, in Abhängigkeit von der Verwendung verschiedener Software-Gruppen. Zu diesem Zweck wählten die Befragten zuvor aus, welche der Software-Gruppen sie überhaupt nutzen. Auf diese Weise stellten wir sicher, dass die Befragten nur angeben konnten, wie unwohl sie sich bei Software fühlen, die sie tatsächlich auch benutzen. Die Ergebnisse hierzu sind in Abbildung 6.8 abgebildet. Die Analyse der Antworten aus (a) resultierte in insgesamt 96 Codes, die wir in neun Kategorien einteilten, in denen sich die Befragten bei der Nutzung der Software unwohl fühlten. Die Kategorien sind in Abbildung 6.9a dargestellt.



(a) Kategorien, in denen sich die Befragten bei der Nutzung der Software unwohl fühlten

(b) Bedrohungen, über die sich die Befragten Sorgen machen

Abbildung 6.9: Wahrgenommene Bedrohungen und Bedenken hinsichtlich der eigenen Privatsphäre

Unsere Daten legten offen, das das größte Unbehagen unter den Befragten durch **exzessive Datenerfassung** ausgelöst wird, gefolgt von **invasiver Werbung**. Invasive Werbung ist dadurch gekennzeichnet, viele verschiedene Daten zu sammeln und zu analysieren, um daraus relevante Werbung auf der Grundlage dieser Daten zu präsentieren. Die Befragten gaben beispielsweise an, dass sie den Eindruck haben, dass Offline-Gespräche von Technologien wie Smarthome-Assistenten aufgezeichnet werden und dass bei der nächsten Online-Suche entsprechende Werbung für entsprechende Produkte angezeigt werde. Andere äußerten ihr Unbehagen über Werbung, die auf dem Chatverlauf von Messaging-Diensten basierte. Des Weiteren berichteten die Teilnehmer, dass sie sich beim **Online Shopping** unwohl fühlen, weil Shopping-Portale häufig viele Daten sammeln, wie Adress- und Zahlungsdaten sowie

Bestellhistorien anlegen und Konsumverhalten dokumentieren. Darüber hinaus beklagten mehrere der Befragten, dass für die Ausführung von Software oftmals viele **Berechtigungen** erforderlich sind (z.B. Zugriff auf die Kamera, Standort usw.), ohne dass klar ersichtlich ist, wofür diese Berechtigungen eigentlich benötigt werden.

Wahrgenommene Bedrohungen Das Ergebnis auf die Frage, von welchem Risiko ihrer Privatsphäre sich die einzelnen Teilnehmern am stärksten bedroht sehen, ist in Abbildung 6.9b dargestellt. Aus den Antworten der Befragten zu ihren Bedenken konnten wir 171 Codes extrahieren. Diese haben wir dann in sieben Kategorien eingeteilt. **Verlust der Kontrolle über die Daten** ist die von den Befragten am häufigsten genannte Sorge. Hier äußerten die Befragten die Sorge, nicht zu wissen, wer Zugang zu ihren Daten hat, mit wem diese Daten geteilt werden und zu welchem Zweck die Daten überhaupt gesammelt werden. Zu den damit zusammenhängenden Bedenken gehören **Datendiebstahl** (*„Ich habe Angst, dass meine Bankdaten gestohlen werden“*, *„Dass eines Tages unsere Identitäten (wie in den Filmen) gestohlen werden können“*), **Datenmissbrauch** (*„Nutzung der Informationen zur Begehung von Straftaten“* und **Profiling** (*„Erstellung und Analyse von eindeutigen Profilen durch Metadatenkonsolidierung“*). Darüber hinaus ist auch die **exzessive Datenerfassung** eine der hier erwähnten Sorgen und einige der Befragten sind besorgt darüber, **ausspioniert zu werden** (*„Software spioniert mich aus“*). In der Kategorie **Sonstiges** wurden Aussagen kategorisiert, die nur aus einem Wort bestanden (z. B. *„E-Mail“*) oder Aussagen, die in keine der anderen Kategorien passen würden.

6.3.3 Aktuelle Wahrnehmung von Privacy Explanations

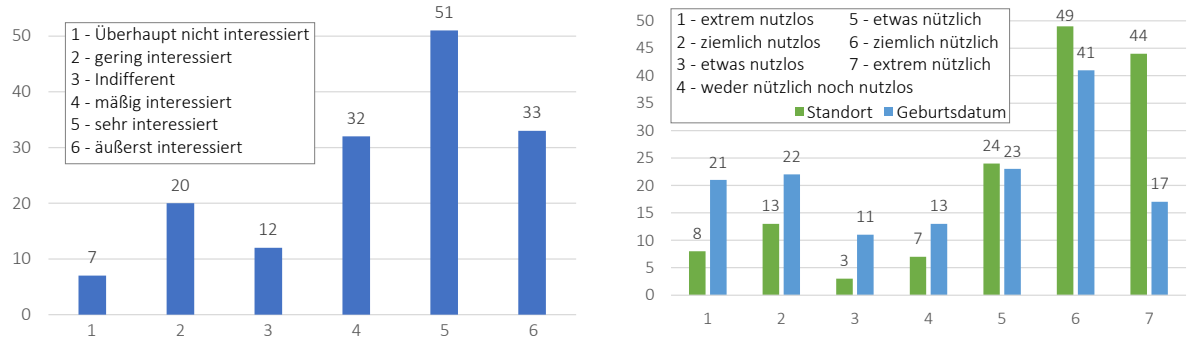
Der vierte und letzte Block der Umfrage (siehe Abbildung 6.2) beschäftigt sich mit den Privacy Explanations. Dafür baten wir die Teilnehmer sich zu Beginn in das nun präsentierte hypothetische Szenario hinein zu versetzen. Ziel war es, das Bedürfnis der Endnutzer nach Privacy Explanations in Situationen zu analysieren, in denen ein Software-System den Nutzer auffordert, persönliche Informationen preiszugeben. Auf der Grundlage dieses Szenarios wurden die Befragten zunächst gefragt, ob sie an einer Erklärung zur Verwendung ihrer persönlichen Daten interessiert seien. Anschließend präsentierten wir den Teilnehmern, die angaben Interesse an einer Privacy Explanation zu haben, jeweils eine Privacy Explanation entsprechend der angeforderten Daten (Privatsphäreaspekt). Die in der Umfrage präsentierten Erklärungen sind in Tabelle 6.3 zu finden.

Insgesamt gaben 148 (95,5%) der 155 Teilnehmer an, dass sie zumindest geringfügig an einer Erklärung interessiert sind. Die genauen Ergebnisse sind in Abbildung 6.10a dargestellt. Den Befragten, die Privacy Explanations wünschten ($n = 148$), wurden anschließend die aus Tabelle 6.3 zu entnehmenden Erklärungen gezeigt. Daraus ging hervor, wie Ihre persönlichen

Tabelle 6.3: Übersicht der gegebenen zwei Privacy Explanations bezüglich des hypothetischen Szenarios

Privatsphäreaspekt	Erklärung
Standort (E1)	Um Ihnen Touren und Empfehlungen in Ihrer Nähe anzeigen zu können, benötigen wir Zugriff auf Ihren Standort.
Geburtsdatum (E2)	Anhand Ihres Geburtsdatums können wir Ihnen Empfehlungen zeigen, was anderen Nutzern Ihres Alters gefallen hat.

Daten in Bezug auf das hypothetische Szenario verwendet werden. Die Rangkorrelation nach Spearman zeigte, dass Befragte mit einem höheren Risk Score statistisch signifikant häufiger an einer Erklärung zum Datenschutz interessiert waren (ρ=0,38, p<0,001). Abschließend wurden die Teilnehmer dann gefragt, wie nützlich sie jede dieser Privacy Explanation fanden (Ergebnisse siehe Abbildung 6.10) und ob die Erklärungen dazu beitrugen, dass sie sich bei der Offenlegung ihrer persönlicher Daten anschließend wohler fühlten, nachdem Sie wussten wofür diese genutzt werden würden.



(a) Inwieweit sind Sie daran interessiert, eine Privacy Explanation zu erhalten? (Hypothetische Situation)

(b) Wie nützlich finden Sie diese Art von Privacy Explanations? (n = 148)

Abbildung 6.10: Interesse an Privacy Explanations und deren empfundener Nutzen

Hierfür gab es drei Antwortmöglkichkeiten. Diese und die Ergebnisse dazu sind in Tabelle 6.4 dargestellt. Befragte, die Antwortmöglichkeit drei wählten, mussten ihre Entscheidung durch Eingabe einer Antwort in ein Textfeld begründen. Die Analyse der Antworten ergab 49 Codes. Wir kategorisierten die Aussagen entsprechend ihrer Bedeutung. Die Kategorien sind **Verbesserungen** (14, 28,6%), **Kritik** (4, 8,2%), **Unzureichend** (14, 28,6%), **Misstrauen** (9, 18,4%) und **Sonstiges** (8, 16,3%).

Tabelle 6.4: Fühlen Sie sich mit den Privacy Explanations wohler? (n = 148)

Antwort	Anzahl
1. Ja, ich fühle mich wohler.	87 (58,8%)
2. Es hat mich vorher schon nicht beunruhigt.	44 (29,7%)
3. Nein, ich fühle mich nicht wohl oder noch nicht ganz wohl, weil:	17 (11,5%)

Die Gruppe **Verbesserungen** umfasste Aussagen wie *„Ich würde die Funktion lieber ausschalten“* in Bezug auf das Geburtsdatum und *„Einteilung in Altersgruppe wäre ausreichend“*. In die Kategorie **Kritik** wurden Aussagen wie *„Interessen hängen nicht vom Alter ab“* aufgenommen. Aussagen wie *„die Argumentation zur Altersfrage ist nicht ausreichend“* sowie *„ich müsste zusätzlich wissen, dass dies die einzigen Gründe sind“* haben wir der Kategorie **Unzureichend** zugeordnet. Die Antworten der Befragten wie *„nicht vertrauenswürdig“* sowie *„fühlen sich ausspioniert“* wurden der Kategorie **Misstrauen** zugeordnet. Aussagen wie *„Die Erklärung ist Unsinn. Ich habe sie nicht gewollt.“* wurden der Kategorie **Sonstiges** zugeordnet. In dem Szenario, das den Befragten vorgelegt wurde, fragte die App nach dem vollständigem Geburtsdatum. Wir haben das Szenario bewusst so konstruiert, dass statt des Geburtsjahres - was für eine Altersempfehlung technisch völlig ausreichend gewesen wäre - das Geburtsdatum abgefragt wurde. Also bewusst detailliertere Daten abgefragt, als notwendig gewesen wären. Sieben der Befragten haben dies explizit erwähnt.

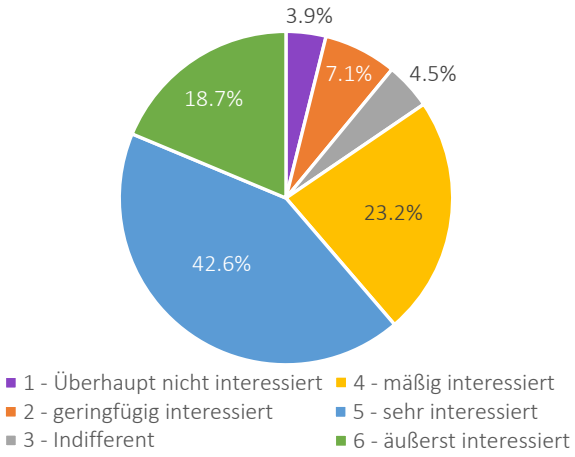
Grundsätzliches Interesse an Privacy Explanations Nachdem sich die Teilnehmer mit Hilfe des hypothetischen Szenarios mit dem Konzept der Privacy Explanations vertraut machen konnten, fragten wir ob generelles Interesse an Privacy Explanations bestünde und wann diese angezeigt werden sollten (siehe Tabelle 6.5). Auf letztgenannte Frage wählten neun Teilnehmer die Antwortoption **Sonstiges**. Zur Begründung mussten Sie ihre Antwort in einem Textfeld formulieren und die Analyse dieser neun Antworten deckte auf, dass die Befragten automatisch eine Erklärung erhalten möchten und zwar jedes Mal wenn sich etwas ändert, aber auch die Möglichkeiten möchten, immer eine Erklärung erhalten, wenn Sie zusätzliche explizit eine anfordern.

Die überwiegende Mehrheit (91,6%) der Teilnehmer gab an, dass sie generell an Erklärungen interessiert sind (*äußerst interessiert* bis *geringfügig interessiert*), wie in Abbildung 6.11a dargestellt. Betrachtet man den Risk Score der Befragten und ihr Interesse an Privacy Explanations, ergibt sich ein ähnliches Bild wie bei der Frage aus Abbildung 6.10a. Es existiert eine positive Korrelation nach Spearman (ρ=0,29, p<0,001). Je höher der Risk Score, desto (statistisch signifikant) höher das Interesse an einer Privacy Explanation.

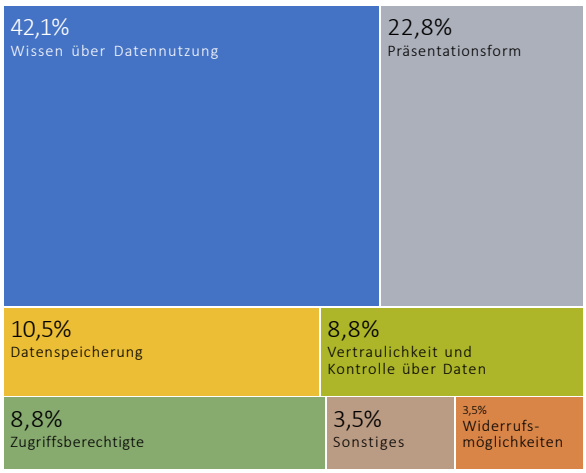
Tabelle 6.5: Wann sollte eine Privacy Explanation angezeigt werden?

Antwort	**Anzahl**
Niemals	2 (1,3%)
Sonstiges	9 (5,8%)
Jedes Mal, wenn ich sie anfordere	39 (25,2%)
Automatisch, wenn etwas in Bezug auf meine Privatsphäre passiert (z.B. Nutzung von persönlichen Daten)	105 (67,7%)

Relevante Aspekte von Privacy Explanations Als Antwort auf unsere offene Frage, was eine Erklärung zum Schutz der Privatsphäre enthalten sollte bzw. was von ihr erwartet wird, ergab die Analyse der Daten 57 Codes, die wir Abbildung 6.11b entsprechend kategorisierten.



(a) Sind Sie generell interessiert an Privacy Explanations?



(b) Wichtige Aspekte, die bei Privacy Explanations berücksichtigt werden sollten

Abbildung 6.11: Interesse an und Aspekte von Privacy Explanations

In erster Instanz wollten die Teilnehmer verstehen, wie ihre Daten verwendet werden. Dazu gehört das Wissen darüber, welche Daten verwendet werden (**Wissen über Datennutzung**), warum und wie. Auch die Art der Darstellung der Erklärungen spielt eine wichtige Rolle (**Präsentationsform**). Die Befragten gaben auch an, dass Erklärungen kurz, präzise und in einfacher Sprache (informell) verfasst sein sollten. Also im vollständigen Kontrast zu Datenschutzerklärungen (DSEs). Darüber hinaus gaben die Befragten zusätzlich an, dass Icons eine visuelle Unterstützung sein können. Teilnehmer äußerten auch, dass sie gerne wissen würden, wo die Daten gespeichert werden, wie lange, wie sie geschützt werden und ob oder wann die Daten gelöscht werden (**Datenspeicherung**). Die Befragten wollten „*sicher sein, dass diese Daten nicht an andere Unternehmen verkauft werden können*“ sowie das die Daten vertraulich behandelt und „*nicht für etwas anderes verwendet werden*“ (**Vertraulichkeit und Kontrolle über Daten**). Eindeutige Informationen darüber wer **Zugriff auf die Daten** hat und zu **Widerrufsmöglichkeiten** ist ebenfalls von Relevanz. Die Kategorie **Sonstiges** umfasste Aussagen, bei denen wir keinen Bezug zu unserer Fragestellung herstellen konnten.

6.3.4 Privacy Explanations und das Konzept von Vertrauen

Der achte und letzte Teil unserer Umfrage beschäftigte sich mit dem Einfluss von Privacy Explanations auf das Vertrauen von Endnutzern gegenüber einem Informationssystem und der Fragen, ob die Teilnehmer einen möglichen Nutzen in Privacy Explanations sehen.

Tabelle 6.6: Was ist der Nutzen von Privacy Explanations?

Code	Häufigkeit	Code	Häufigkeit
Transparenz	39,4%	Vertrauensfördernd	12,9%
Wohlbefinden	11,0%	Rechtliche Möglichkeiten	6,9%
Sonstiges	6,9%	Conscious Choices	5,5%
Fördert Privacy Awareness	4,1%	Selbstbestimmung	2,8%
Umdenken bei Datenpraktiken	1,7%	Augenwischerei	1,7%
User eXperience	1,7%	Geringer Vorteil	1,4%
Positive Company Image	1,4%	Kundenbindung	1,1%
Absicherung	0,8%	Misstrauen	0,8%

Möglicher Nutzen von Privacy Explanations Die optionale offene Frage nach dem möglichen Nutzen von Privacy Explanations wurde von 137 Teilnehmern beantwortet. Davon konnten wir 135 Antworten als valide betrachten. Die Analyse dieser 135 validen Antworten resultierte in 363 Codes, die wir in 16 Kategorien einordneten. Das Ergebnis und die prozentuale Verteilung dieser 16 Kategorien ist in Tabelle 6.6 abgebildet. In der Kategorie **Sonstiges** haben wir 24 Codes (6,6%) zusammengefasst, die wir keinem Vorteil zuordnen konnten. Diese Aussagen hätten auch nicht einem Nachteil zugeordnet werden können, da sie teilweise auf unzureichendes Verständnis seitens der Befragten schließen ließen. Wir gruppierten Aussagen wie zum Beispiel *„Datenverkauf“* oder *„minimiert Spam“* hier ein. Da wir in solche Aussagen zu viel hätten hinein interpretieren müssen, wäre eine Einordnung in eine andere Kategorie zu subjektiv. Daher haben wir uns entschieden, solche Aussagen in diese Kategorie einzuordnen. Ein Überblick mit exemplarischen Auszügen von Aussagen der Befragten ist zusätzlich in Tabelle 6.7 zu finden.

Tabelle 6.7: Auszüge von Aussagen der Teilnehmer bezüglich dem Nutzen von Privacy Explanations

Kategorie (Anzahl an Codes)	Aussagen der Teilnehmer
Absicherung (3)	„Schutz der eigenen Interessen“ „Schutz der eigenen Person vor z.B. unnötiger Werbung“
Augenwischerei (6)	„den Nutzern was vormachen“ „Verwirrung der Verbraucher (Verharmlosung)“
Conscious Choices (20)	„Bedarfsermittlung“ „gibt mit mehr Wahlmöglichkeiten“
Fördert Privacy Awareness (15)	„Sinn für Privatsphäre“ „man wird mehr sensibilisiert“
geringer Vorteil (5)	„Nutzen gering“ „Nutzern gering, weil oft nicht actionable“
Kundenbindung (4)	„Steigerung der Akzeptanz beim Endkunden“ „Ermutigung, eine vertrauenswürdige Anwendung an eine andere Person weiterzuleiten“
Misstrauen (3)	„man kann nicht darauf vertrauen, dass es eingehalten wird“ „vielleicht wird nur die Hälfte der Nutzung verraten“
Positive Company Image (5)	„Vertrauen zum Anbieter“ „Glaubwürdigkeit des Unternehmens“
rechtliche Möglichkeiten (25)	„Haftbarkeit im Fall von Missachtung“ „Nutzung von Daten juristisch regeln“
Selbstbestimmung (10)	„Kontrolle über die eigenen Daten“ „Entscheidung darüber, ob man das möchte“
Sonstiges (25)	„Datenverkauf“ „das Unternehmen verdient damit Geld“
Transparenz (143)	„erhöht Verständnis der Benutzer für die Software, die sie benutzen“ „die Möglichkeit, Dienstleistungen zu vergleichen“
Umdenken bei Datenpraktiken (6)	„der Hersteller wird gezwungen, sich über die Verwendung der Daten Gedanken zu machen“ „Als Entwickler der Verwendung und Notwendigkeit von Daten bewusst werden“
User Experience (6)	„Benutzerfreundlichkeit“ „Bessere User Experience“
Vertrauensfördernd (47)	„Aufbau von Vertrauen gegenüber dem System“ „... was schlussendlich das Vertrauen erhöhen kann“
Wohlbefinden (40)	„Beruhigung der Nutzer“ „Ich fühle mich sicher, wenn ich weiß, wie meine Daten behandelt/-verwendet werden“

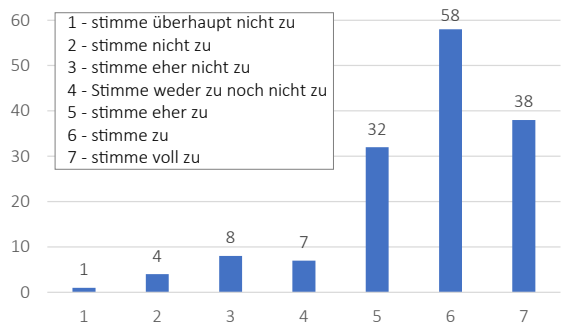


Abbildung 6.12: Können Privacy Explanations ein möglicher Faktor sein, um das Vertrauen in ein Software-System zu erhöhen? (n = 148)

Privacy Explanations und Vertrauen Der Frage, ob *Privacy Explanations ein möglicher Faktor sein könnten, um das Vertrauen in ein Software-System zu erhöhen* stimmten 86,5%[6] der Teilnehmer zu (siehe Abbildung 6.12). Hierbei wurden nur die Teilnehmer gefragt, die beim hypothetischen Szenario an einer Privacy Explanation interessiert waren (n = 148). Die Antworten der Teilnehmer deckten sich hierbei, mit der Ergebnissen aus Tabelle 6.6, in den als wichtiger Nutzen von Privacy Explanations Vertrauen auch explizit genannt wurde.

6.4 Einschränkungen und Bedrohung der Validität

Die Strategie zur Auswahl der Teilnehmer weist einige Einschränkungen auf. Trotz der Tatsache, dass wir Antworten aus verschiedenen Ländern der Welt erhielten, kam die Mehrheit der Antworten aus Brasilien und Deutschland. Dies spiegelt möglicherweise nicht die gesamte Population wider und könnte die generelle Verallgemeinerbarkeit unserer Ergebnisse gefährden. Obwohl 155 Teilnehmer eine beträchtliche Anzahl von Antworten lieferten, sollten einige der Schlussfolgerungen nicht zu stark verallgemeinert werden, da diese durch die Anzahl der Teilnehmer beeinflusst sein könnten. Die meisten Befragten unserer Studie verfügten über fundierte IT-Kenntnisse. In unserer Population sind Personen, die Schwierigkeiten mit der Bedienung von Softwaresystemen hatten, möglicherweise nicht berücksichtigt. Daher können wir den Bedarf an Erklärungen zum Schutz der Privatsphäre nicht verallgemeinern, aber wir erhalten einen validen Überblick darüber, was verschiedene Personen denken. Für Q2 ermitteln wir unterschiedliche Bedenken bei der Nutzung von Software. Wir können nur die Antworten unserer Befragten auswerten und es könnte sein, dass es viel mehr Gründe für Bedenken bezüglich der Privatsphäre gibt, z.B. bei Personen, die keine IT-Kenntnisse haben. Die anderen Ergebnisse unserer Untersuchung weisen dieselben Einschränkungen auf. Um weitere Bedenken im Bereich der Privatsphäre zu ermitteln, müssten weitere Experimente durchgeführt werden. Um das Problem, dass die Analyse zu subjektiv ist, zu entschärfen, verwendeten wir *In Vivo Coding*, das von zwei Forschern unabhängig voneinander durchgeführt wurde. In einem zweiten Kodierungszyklus diskutieren und vergleichen wir unsere Ergebnisse, um die Konsistenz und Zuverlässigkeit zu erhöhen.

[6] Aggregierung der Antworten von *stimme voll zu*, *stimme zu* und *stimme eher zu*

Um den Bedarf an Privacy Explanations zu bewerten, verwenden wir ein hypothetisches Szenario. Dieses Szenario ist für die Teilnehmer möglicherweise nicht alltäglich, aber es könnte ihnen im wirklichen Leben begegnen. Allerdings konnten sich Nutzer von Smartphones möglicherweise besser in diese Situation hineinversetzen. Das Szenario konfrontierte die Befragten nur in einem Urlaubsszenario. Die Ergebnisse könnten anders ausfallen, wenn sie mit einem Geschäfts- oder Finanzszenario konfrontiert worden wären, da dies möglicherweise eine größere Bedrohung für ihre Privatsphäre hätte darstellen können.

Ein weiterer Aspekt ist, dass eine gute Formulierung der Fragen und die Survey-Gestaltung für die Ergebnisse einer Umfrage entscheidend sind. Wir haben uns an Leitlinien gehalten und Pilotierungen durchgeführt, um diese Aspekte sicherzustellen. Die Reihenfolge der Fragen im Fragebogen könnte sich jedoch auf das Verständnis der Teilnehmer ausgewirkt haben, ob wir Fragen zum Erklärungsbedarf in einem allgemeinen Kontext oder im Zusammenhang mit der vorherigen Frage zu einem spezifischeren Kontext stellten. Wir waren jedoch der Ansicht, dass dies für die Teilnehmer hilfreich sein würde, die sich möglicherweise nur schwer andere Situationen vorstellen können, in denen sie Erklärungen benötigen.

Wir haben beschlossen, unseren Online-Fragebogen und die Rohdaten offenzulegen [386], damit andere Forscher nachvollziehen können, wie wir unsere Schlussfolgerungen und Empfehlungen aus den Daten gezogen haben. Dieser Schritt sollte als letzte Strategie zur Abschwächung der Gefahren für die interne Validität dienen.

6.5 Fazit

Laut unseren Ergebnissen ist die Mehrheit (91,6%) der Befragten grundsätzlich an Privacy Explanations interessiert. Ein genauerer Blick auf die im hypothetischen Szenario gegebenen Erklärungen und die Reaktionen der Befragten auf diese zeigen, dass die Teilnehmer die Privacy Explanations als unterstützend wahrnahmen und sich wohler fühlten (Tabelle 6.4), wenn sie Informationen zu etwaigen Datenpraktiken erhielten. Durchaus interessant ist zudem die Beobachtung, dass das bloße Vorhandensein einer Erklärung scheinbar nicht ausreicht, um einen positiven Effekt zu erzielen. Das zeigte der unterschiedlich wahrgenommene Nutzen der Erklärungen E1 und E2 (Tabelle 6.3). 74,3% empfanden E1 als nützlich und waren damit zufrieden, aber nur 58,8% bestätigten dies für E2. Den Grund dafür haben unsere Teilnehmer auch genannt (siehe Abschnitt 6.3.3), denn es hätte technisch gesehen vollkommen ausgereicht, nur das Geburtsjahr abzufragen.

Den Sinn und Zweck und somit den Nutzen, den Privacy Explanations für die Befragten zu erfüllen scheinen, konnten wir ebenfalls im Rahmen der Studie erfolgreich untersuchen.

Die Befragten gaben an, dass Erklärungen ihre Unsicherheit verringern, für mehr Sicherheit bei ihnen sorgen und Vertrauen vermitteln, was wiederum zu einem allgemeinen Gefühl des Wohlbefindens führt. Des Weiteren deuten unsere Ergebnisse darauf hin, dass die Informierung und Aufklärung eines Nutzers darüber, dass personenbezogene Daten über ihn gesammelt werden, welche Daten gesammelt werden und wie diese Daten verwendet werden, zur Sensibilisierung für den Datenschutz (Privacy Awareness) beitragen. Dieses Phänomen wird ebenfalls von Poetzsch [114] beschrieben und bestätigt. Ergänzend dazu ermöglicht das *informiert sein* selbstbestimmte und bewusste Entscheidungen bei der Nutzung von Softwaresystemen (explizite Zustimmung), was ebenfalls von den Befragten erwähnt wurde. Folgt man diesem Gedankengang, wird auch klar, warum einige der Befragten Privacy Explanations als eine Art *Safeguard* betrachten. Also ein Schutz oder eine Absicherung der eigenen Online-Privatsphäre und somit ein Schutz vor der uninformierten, ungewollten Offenlegung von persönlichen Informationen. Ein weiterer Blick auf unsere Ergebnisse zeigt, dass Privacy Explanations ein Mittel sein können, um das Vertrauen in Software-Systeme zu erhöhen (Abbildung 6.12). Privacy Explanations könnten also dazu beitragen, ein Vertrauensverhältnis zwischen Endnutzer und System aufzubauen, indem sie für Transparenz und Klarheit im Umgang mit den persönlichen Daten sorgen. Unter Berücksichtigung der Ergebnisse scheinen Privacy Explanations eine sinnvolle und vor allem nutzerzentrierte Lösung darzustellen, die den Datenschutz *usable* und *actionable* macht.

Anforderungen an Privacy Explanations Um Privacy Explanations in Informationssysteme einbauen zu können, müssen die Anforderungen der relevanten Stakeholder elizitiert werden. Es ist wichtig, die Bedürfnisse und Erwartungen der Stakeholder in Bezug auf Erklärungen zu erfüllen, andernfalls könnten sie ihren Zweck, den Endnutzer über seine Privatsphäre zu informieren, verfehlen oder sogar Misstrauen hervorrufen [164, 387]. Das zeigten unsere Ergebnisse ebenfalls. In Anbetracht dessen reicht es also nicht aus, „beliebige“ Erklärungen zu liefern. Eine Privacy Explanation muss für den Endnutzer unter Berücksichtigung seiner individuellen Vorlieben und seines Kontexts sinnvoll und nachvollziehbar sein.

Dies liefert Software Engineers wichtige Erkenntnisse für die Entwicklung erklärbarer Systeme, die das Qualitätsziel verfolgen, *privacy-aware* zu sein. Endnutzer fordern eine Reduzierung des Datensammelns und Datensparsamkeit, weil ihnen, wie oben erwähnt, ihre Privatsphäre wichtig ist. Das bedeutet, dass für das Design von Informationssystemen ein sparsamer, verantwortungsvoller und fairer Umgang mit personenbezogenen Daten erforderlich ist [388, 389]. Für unser Szenario bedeutet dies, dass dieses System statt nach dem Geburtsdatum nur nach dem Geburtsjahr fragen sollte. Darüber hinaus und unter Berücksichtigung der Verbesserungsvorschläge der Befragten sollte eine Erklärung in der Lage sein, bei Bedarf weitere Informationen zu liefern. Zum Beispiel, welche Folgen es für den Nutzer hat, wenn er

seine Zustimmung bei der Datenverarbeitung nicht gibt. Eine Privacy Explanation sollte also mehrere *Layer* (Ebenen) enthalten, mit denen ein Endnutzer die Granularität der gewünschten Informationen selbstständig, entsprechend seinen Bedürfnissen, bestimmen kann.

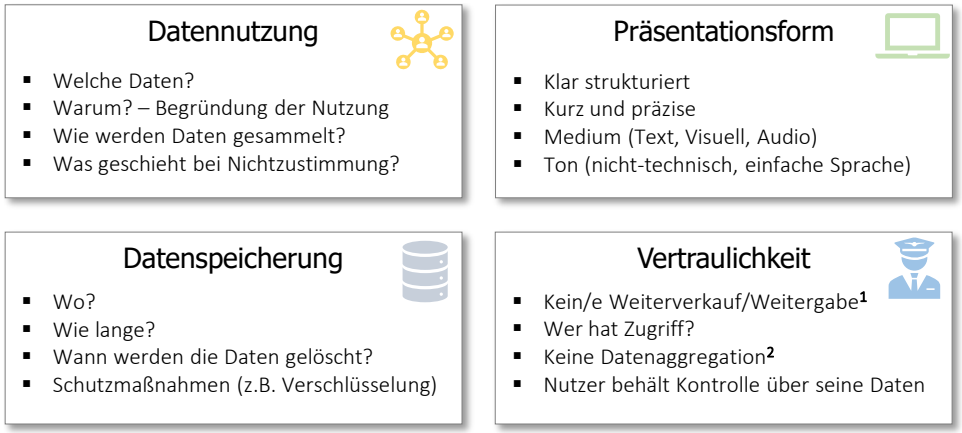


Abbildung 6.13: *High-Level* Anforderungen an Privacy Explanations

Wir haben die Teilnehmer gefragt, was sie von Privacy Explanations erwarten bzw. was diese enthalten sollten (Abbildung 6.11b). Dies half uns in Übereinstimmung mit unserer Definition von Online-Privatsphäre (Abschnitt 6.1.2), Aspekte zu identifizieren, die in Privacy Explanations berücksichtigt werden sollten. Diese Aspekte können als high-level Anforderungen an Privacy Explanations (siehe Abbildung 6.13) behilflich sein. Die dann wiederum als Ausgangspunkt für Software Engineers dienen können, um zu verstehen, welche Elemente beim Entwurf von Privacy Explanations berücksichtigt werden sollten. Für eine Diskussion dieser Anforderungen verweise ich an dieser Stelle auf Kapitel 7. Mit Hilfe der aus unseren Ergebnissen gewonnen Erkenntnissen kann ich nun die Antwort für RQ2 wie folgt formulieren:

Beantwortung RQ2: Das Konzept der Erklärbarkeit, kombiniert mit den Prinzipien des PbD-Konzepts und minimalen (Informations-) Asymmetrie, bilden die gemeinsame Basis für *Privacy Explanations*. Mit Hilfe von Privacy Explanations können Antworten und Hinweise zur Datennutzung an den Endbenutzer kommuniziert werden und dadurch möglichen Bedenken entgegengewirkt, Vertrauen aufgebaut, Transparenz geschaffen und eine Sensibilisierung für den Schutz der persönlichen Daten erreicht werden, die den Benutzer in die Lage der expliziten Zustimmung versetzt. Darüber hinaus stellen die Anforderungen an Privacy Explanations (Abbildung 6.13) ein weiteres wichtiges Artefakt für Software Engineers bereit, wie diese in auf den Datenschutz ausgerichteten Systemen integriert werden können.

7

Vom Konzept hin zur technischen Umsetzung von Privacy Explanations

Im vorherigen Kapitel 6 wurde untersucht, wie das Konzept der Erklärbarkeit eingesetzt werden kann, um Endbenutzern die Verwendung von Privatsphäreaspekten auf verständliche Art und Weise zu kommunizieren. Hierzu wurde das Artefakt der *Privacy Explanation* geschaffen. Die Daten der Umfrage haben hierbei wertvolle Einblicke geliefert wie Endbenutzer Privacy Explanations wahrnehmen und welche unterschiedlichen Bedarfe sie an diese haben. In diesem Kapitel soll dieses geschaffene Wissen nun validiert und vertieft werden sowie Konzepte und Leitlinien vorgestellt werden, die für die technische Umsetzung von Privacy Explanations notwendig sind.

Zugehörige Publikationen Der in diesem Kapitel vorgestellte Forschungsbeitrag entstand in Kollaboration[1] mit zwei anderen Forschern: Jakob Droste und Kurt Schneider. Dieses Kapitel gründet auf der gemeinsamen Zusammenarbeit und zwei weiteren Masterarbeiten von Jakob Droste [5] und Felix Volodarskis [6], die von mir betreut wurden. Darüber hinaus wurden einige der Ergebnisse hierzu in [390] veröffentlicht.

[1] Aufgrund der Kollaboration mit meinen Forschungskollegen verwende ich das „wir“.

7.1 Forschungsvorgehen

Die Forschungen zu diesem Kapitel bauen auf den gewonnen Daten und Erkenntnissen der vorangegangenen Kapiteln (maßgeblich Kapitel 6) auf und stellen somit die stringente Weiterentwicklung des Konzepts der Privacy Explanations dar. Hierbei wurde die Umsetzung der Privacy Explanations zunächst in einer Konzeptstudie evaluiert, um somit die Ergebnisse und Daten in die Verfeinerung und Verbesserung für einen technischen Prototypen einfließen zu lassen. Der technische Prototyp wurde anschließend ebenfalls in einer Benutzerstudie evaluiert. Abbildung 7.1 veranschaulicht hierzu das gesamtheitliche Forschungsvorgehen.

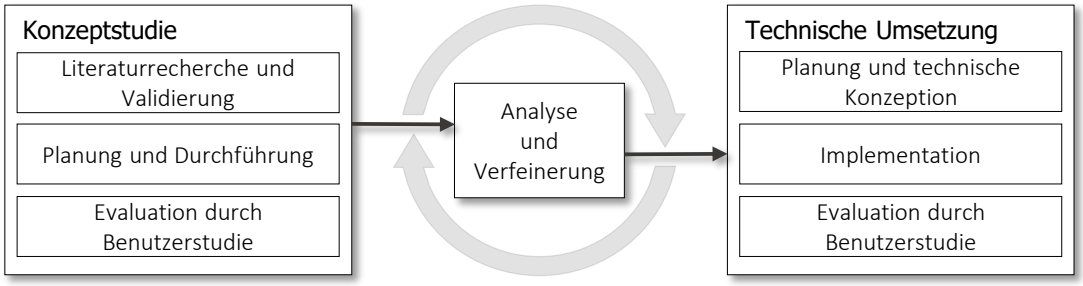


Abbildung 7.1: Überblick des Forschungsvorgehens

Im Vorfeld der Konzeptstudie haben wir eine LR mit dem Ziel durchgeführt, Erkenntnisse und Daten - an erster Stelle die High-Level Requirements an Privacy Explanations - aus unserer Umfrage (Kapitel 6) zu validieren und tiefere Erkenntnisse zu gewinnen. Aufbauend auf diesen Daten haben wir für unsere Konzeptstudie einen *high-fidelity* Prototyp entworfen und diesen nachfolgend in einer Benutzerstudie evaluiert. Bezugnehmend auf Jacobsen und Meyer [391] ist ein high-fidelity Prototyp ähnlich zu einem Mockup, beinhaltet jedoch interaktive Elemente und hat bereits ein dem finalen Produkt ähnelndes *look and feel*. Details zur durchgeführten LR, sind nachfolgend in Abschnitt 7.1.1 und zur Datenvalidierung in Abschnitt 7.1.2 zu finden. Mit Hilfe der erhobenen Daten aus der Benutzerstudie unseres high-fidelity Prototypen haben wir das Konzept der Privacy Explanations verfeinert und anschließend die technische Umsetzung begonnen. Details zur technischen Umsetzung werden in Abschnitt 7.3 diskutiert.

7.1.1 Literaturrecherche

Ein Überblick über den angewandten Prozess unserer LR ist in Abbildung 7.2 dargestellt. Spezifische Details zum Prozess wie Einschluss- und Ausschlusskriterien, identifizierte Literatur oder auch die Such-Strings für die Datenbanksuche sind in Abschnitt E.1 zu finden. Den Ausgangspunkt für die LR bildete die Literatur, die wir in [4] (*Baseline Paper*) in Zusammenhang mit Erklärbarkeit und Privatsphäre identifiziert haben. Auf Basis einer Relevanzprüfung

dieser Arbeiten erfolgte die Auswahl unseres *Startsets*, welches wir anschließend mit Snowballing komplementierten. Eine Publikation wurde als relevant identifiziert, wenn Sie unseren Einschluss- und Ausschlusskriterien (Abschnitt E.1) entsprach bzw. nicht gegen diese verstieß. Unser Startset hatte einen Umfang von 42 Publikationen. Beim Snowballing folgten wir den Richtlinien von Wohlin [225].

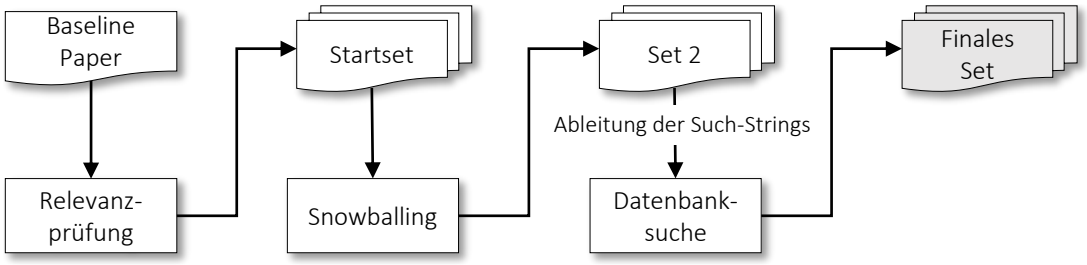


Abbildung 7.2: Prozess der Literaturrecherche

Das Snowballing mündete in weiteren 45 als relevant erachteten Arbeiten und bildete das *Set 2*. Der Fokus der Inhalte dieser insgesamt 87 Publikationen (Startset und Set 2) lag primär auf Erklärbarkeit als NFR und XAI. Um einen zusätzlich stärkeren Fokus auf Privatsphäre zu legen, komplementierten wir unsere Suche mit einer Datenbanksuche. Diese lieferte schließlich 42 weitere Publikationen. Unser LR-Prozess basierte ebenfalls auf dem von Wolfswinkel et al. [229] vorgeschlagenen Ansatz der GT für Literaturrecherchen, sodass bereits eine erste Iteration eines Snowballings nach der Datenbanksuche keine neuen Erkenntnisse lieferte.

7.1.2 Analyse und Validierung

Bei der Analyse lag der Fokus darauf herauszufinden, was die Literatur für Empfehlungen bei Erklärungen in Informationssystemen gibt und wie diese mit Schwerpunkt auf Privatsphäre umzusetzen sind unter Berücksichtigung der Ergebnisse aus Abschnitt 4.3, was das Zusammenspiel von Erklärbarkeit und Privatsphäre angeht. Darüber hinaus haben wir unsere High-Level Anforderungen, die wir auf Basis einer Umfrage erhoben haben (Absatz 6.5) mit Hilfe der Literatur validiert. Hierbei haben wir geprüft, ob die Literatur konkrete Handlungsempfehlungen gibt, was bei Erklärungen an die Privatsphäre berücksichtigt werden sollte oder ob wir auf Basis der gemachten Aussagen Handlungsempfehlungen ableiten können. Diese haben wir anschließend mit den Anforderungen aus Absatz 6.5 verglichen.

7.2 Konzeptstudie und Evaluation

Die gewonnen Erkenntnisse und gesammelten Daten haben wir genutzt, um ein erstes prototypisches Konzept für Privacy Explanations umzusetzen, welches wir anschließend in einer

Benutzerstudie evaluiert haben. Nachfolgend stelle ich zunächst die Ergebnisse der Analyse vor und beschreibe dann die Umsetzung des Prototyps sowie dessen Evaluation.

7.2.1 Einflussfaktoren

Um ein tieferes Verständnis für die Umsetzung von Privacy Explanations zu erlangen ist es nötig, die verschiedenen Faktoren zu identifizieren, die Privacy Explanations beeinflussen. Abbildung 7.3 bildet die von uns identifizierten Einflussfaktoren ab. Diese sind die Synthese aus den Ergebnissen der Umfrage zu Privacy Explanations (Kapitel 6) und unserer LR.

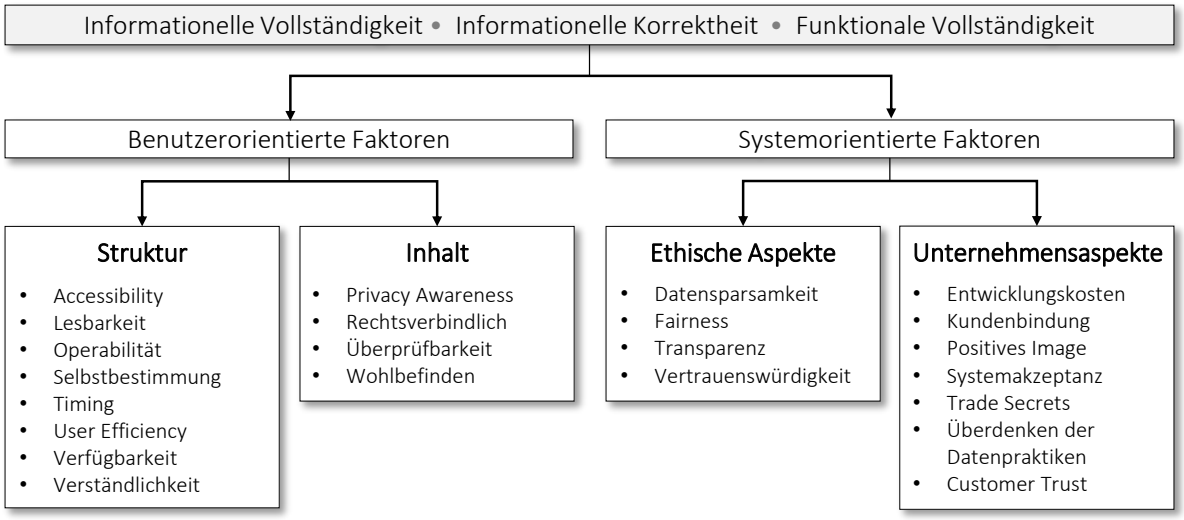


Abbildung 7.3: Einflussfaktoren von Privacy Explanations

Faktoren wie **informationelle** und **funktionale Vollständigkeit** sowie **informationelle Korrektheit** sind als übergelagerte Faktoren zu betrachten, die sowohl die Benutzer- als auch Systemebene betreffen. Diese richten sind an allgemeine Qualitätsaspekte, die Privacy Explanations erfüllen müssen. Um geltenden Gesetzen und Rechtsvorschriften zu entsprechen, müssen Privacy Explanations sowohl vollständig als auch korrekt bzgl. ihrer zu transportierenden Information sein. Das bedeutet, wenn ein Informationssystem beispielsweise Standortdaten einen Benutzers verwendet, muss dieses dem Benutzer nicht nur mitgeteilt werden, sondern das System muss auch offenlegen, wofür diese Daten verwendet werden. Der Benutzer muss also erfahren, *dass* Daten verwendet werden, *wofür* diese Daten genutzt werden und das System muss *garantieren*, dass diese Daten ausschließlich zu diesem Zweck genutzt werden. Das System gibt also eine rechtlich bindende Garantie, die gleichzeitig den Systementwicklern als Nachweis der Rechtskonformität dient.

Damit die in den benutzerorientierten Faktoren beschriebene Funktionsfähigkeit von Privacy Explanations und die allgemeine Nutzbarkeit des Systems gewährleistet werden

kann, muss die Schnittstelle, die die Privacy Explanations bereitstellt, **funktional vollständig** sein. Die funktionale Vollständigkeit betrifft also sowohl benutzerorientierte als auch systemorientierte Faktoren.

Benutzerorientierte Faktoren Die benutzerorientierten Faktoren sind unterteilt in Einflussfaktoren, die die Struktur sowie den Inhalt betreffen und spiegeln somit Bedarfe der Benutzer wider. Damit ein Endbenutzer gemäß den Vorgaben der DSGVO überhaupt zum Thema Datenverarbeitung und Datenschutz informiert werden kann, müssen diese Informationen überhaupt erst einmal **verfügbar** und zugreifbar (**accessibility**) sein. Sunyaev et al. [392] sowie auch Robillard et al. [185] geben zu Bedenken, das Smartphone Apps, allem voran Gesundheit-Apps häufig gar keine derartigen Informationen bereitstellen [392] und längst nicht jede Internetseite stellt Informationen zur Datennutzung bereit [45]. Aber auch wenn diese Informationen vom Anbieter zur Verfügung gestellt werden, sind diese in den meisten Fällen nicht leicht zugänglich [45, 48]. Wichtig für Privacy Explanations ist das **Timing**, da es sich sowohl auf die **Verständlichkeit**, die **User Efficiency**, **Selbstbestimmung** und **Operabilität** auswirkt [393]. Die richtigen Informationen zum richtigen Zeitpunkt präsentiert, machen es dem Benutzer einfacher zu verstehen, warum gerade diese Art persönlicher Informationen benötigt werden und lassen ihn bei gegebener Operabilität aktiv die Entscheidung treffen, ob er der Nutzung seiner Daten zustimmt oder diese ablehnt. Dieses steigert deutlich die geforderte User Efficiency [46, 268, 282, 394], wie sie bei regulären DSEs fehlt.

Bezüglich des Inhalts von Privacy Explanations lassen sich verschiedene Ziele adressieren, die die Bedarfe der Benutzer widerspiegeln. Die Primärziele von Privacy Explanation sind Endbenutzer über Datenpraktiken eines digitalen Informationssystems zu informieren und aufzuklären, um somit ihre **Privacy Awareness** zur erhöhen. Privacy Awareness ist von entscheidender Bedeutung für die Förderung eines sicheren Verhaltens in Bezug auf den Schutz der Privatsphäre und die Überwindung des Datenschutzparadoxons [41, 111, 114, 395]. Darüber hinaus fördern Privacy Explanations durch ihre verständliche und aufklärende Rolle das Wohlbefinden von Nutzern [41], denn „die ordnungsgemäße Information der Nutzer über den Zweck des Ressourcenzugangs kann die Bedenken der Nutzer hinsichtlich des Schutzes ihrer Privatsphäre bis zu einem gewissen Grad verringern“, bestätigen Lin et al. [396]. Des Weiteren zeigen unsere Umfrageergebnisse (Absatz 6.3.4), dass Nutzer eine gewisse **Rechtsverbindlichkeit** in Privacy Explanations sehen bzw. sich diese wünschen in Bezug auf Ihre Privatsphäre. Die Herausforderung hierbei ist, auch wenn die gegebene Erklärung von technischer Seite her korrekt und vollständig in ihrer Aussage ist, mangelt es an Mechanismen der **Überprüfbarkeit** auf Seiten der Endbenutzer. Diesem Gedankengang folgend, ist also eine Form der Verifizierbarkeit der gegebenen Erklärung durch den Endbenutzer notwendig, um das Vertrauen in eine Privacy Explanation als gerechtfertigt anzusehen.

Systemorientierte Faktoren Die systemorientierten Faktoren umfassen die Aspekte, die das System selbst betreffen, wie auch die übergeordneten Business Goals eines Unternehmens, welches das entsprechende System bereitstellt. **Ethische Aspekte** spielen bei Privacy Explanations eine wichtige Rolle. Hier sind natürlich ein **fairer** [184] und **transparenter** [189] Umgang mit den Daten zu nennen sowie die Nutzung von persönlichen Daten zu begrenzen, um **Datensparsamkeit** zu erreichen. Hier sollte das Prinzip der minimalen Informationsasymmetrie greifen. Der Anbieter sollte demnach nur ein Minimum an Informationen verlangen, „so dass nur so viele personenbezogene Daten verarbeitet werden, wie erforderlich sind, erklärt werden und für die eine Einwilligung vorliegt“[2]. Das System selber bzw. deren Designer müssen die **Vertrauenswürdigkeit** der gegebenen Erklärung garantieren. Sie müssen also sowohl informationelle Vollständigkeit als auch Korrektheit liefern. Das heißt, dass das System die in der Erklärung angeführten Privatsphäreaspekte ausschließlich für den beschriebenen Zweck verwenden darf.

Unternehmensaspekte und Privacy Explanations beeinflussen sich ebenfalls. Der empfundene Mehrwert eines Systems kann positiv durch das Vorhandensein von Erklärungen beeinflusst werden [397], weshalb Privacy Explanations ebenfalls direkten Einfluss auf die **Systemakzeptanz** ausüben und diese steigern können [254]. Ist ein Benutzer zufrieden mit einem System und im Kontext der Privacy Explanations informiert und aufgeklärt, wirkt sich das positiv auf sein empfundenes Image eines Unternehmens aus. Es kann zu einer stärkeren **Kundenbindung** beitragen [288] und den **Customer Trust** stärken [281]. Privacy Explanations erhöhen die Transparenz eines Systems, können aber wie auch Erklärungen im Allgemeinen **Trade Secrets** offenbaren und daher die Sicherheit sensibler Unternehmensinformationen gefährden [398]. Auf der anderen Seite, können Erklärungen zur Privatsphäre aber auch einen positiven Einfluss auf das **Überdenken der Datenpraktiken** haben. Greift ein System z.B. möglichst viele private Informationen seiner Nutzer ab, nur für den Fall, diese eventuell mal nutzen zu können, muss es natürlich den Grund in der gegebenen Erklärung angeben. Ist hier kein plausibler Grund sowohl für den Benutzer als auch möglicherweise für das Unternehmen zu erkennen, kann das zu einem sparsameren Umgang mit den privaten Daten der Benutzer führen. Erklärungen zum Datenschutz zu geben erfordern Mehraufwand bei der Entwicklung und resultieren somit in höheren **Entwicklungskosten**. Die Erklärungen sollten allerdings als Chance gesehen werden, da eine Reihe von Aspekten wie ein positives Firmen-Image, Kundenbindung, Systemakzeptanz oder auch Kundenvertrauen erreicht werden können, die im Umkehrschluss durchaus das ökonomische Potential besitzen, die Mehrkosten zu relativieren.

Die in Abbildung 7.3 aufgeführten Einflussfaktoren haben nicht den Anspruch vollständig zu sein. Das Modell dient primär einer ersten Strukturierung des erlangten Wissens durch die

[2] `https://privacypatterns.org/patterns/Minimal-Information-Asymmetry.html#solution`, Letzter Zugriff: 22.09.2023

Auswertung der Umfragedaten sowie der durchgeführten LR und soll beim systematischen Verständnis bis hin zur technischen Umsetzung von Privacy Explanations unterstützen.

7.2.2 Das 3C-Prinzip — Context, Content, Consent

Die eingehende Analyse der Literatur hat ergeben, dass Privacy Explanations sich im Wesentlichen in drei Kernpunkten einbetten lassen: **Context**, **Content** und *informed* **Consent**. Im Weiteren bezeichne ich dies als das **3C-Prinzip**. Im Allgemeinen ist ein Kontext charakterisiert durch eine bestimmte Situation, bei der eine Person, ein System, eine Aufgabe sowie eine Umgebung, in der es stattfindet, involviert sind [242]. Im Fall von Privacy Explanations sprechen wir von **Context**, wenn ein System Privatsphäreaspekte verarbeitet und/oder anfordert oder explizit oder implizit Berechtigungen benötigt (z.B. der Zugriff auf eine Webcam), um eine bestimmte Aufgabe durchzuführen. Mit Verweis auf unsere Definition von Privacy Explanations (Abschnitt 6.1.3) ist diese immer in einen Kontext eingebettet (kontextabhängig) [214, 42, 192]. Daher sollte ein System den Benutzer immer kontextabhängig mit den erforderlichen Informationen versorgen – der Privacy Explanation. Shulman et al. argumentieren ebenfalls, dass Privacy Explanations „nur sinnvoll sind, wenn diese relevant zur momentanen Aktion des Benutzers sind“ [192].

Der **Content** einer Privacy Explanation steht immer im Einklang mit dem Context und bezieht sich daher direkt auf diesen. Der Content spielt eine entscheidende Rolle in Bezug auf die Effektivität der gegebenen Information [192, 399] und wird durch das Design und die Struktur der präsentierten Erklärung gestaltet [184, 189]. Das impliziert, das der Content einer Privacy Explanation sich ebenfalls an Bedürfnissen des jeweiligen Endbenutzers orientiert [254].

Wenn nun das System in einem gegebenen Context, dem Benutzer eine Privacy Explanation präsentiert, geschieht dies mit dem Ziel, den Benutzer über Art und Weise der Verwendung eines spezifischen Privatsphäreaspekts zu informieren und aufzuklären. Hierdurch soll der Benutzer nun in die Lage versetzt werden, bewusste Entscheidungen zu treffen sowie die Kontrolle über seine eigenen Daten zu behalten [268, 282]. Dem Benutzer soll es also möglich sein, seine explizite Zustimmung (engl.: informed **consent**) oder Ablehnung zu Nutzung seiner persönlichen Daten zu geben [279, 394, 191].

7.2.3 Anforderungen an Privacy Explanations

Auf Grundlage der Umfragedaten aus Kapitel 6, war es möglich Anforderungen an Privacy Explanations zu extrahieren, die den Bedürfnissen und Erwartungen von Endbenutzern entsprechen. Mit Hilfe der durchgeführten LR konnten diese Anforderungen validiert werden.

Hierzu haben wir die identifizierte Literatur dahingehend untersucht, was die jeweiligen Autoren empfehlen und vorschlagen hinsichtlich eines transparenten Dialogs zwischen Benutzer und System in Bezug auf die Verwendung privater Informationen bzw. was an Mangel bisheriger Kommunikation in der Literatur festgestellt wurde. Darüber hinaus dienten die DSGVO, der CPRA, die FIPPs und die ISO 29100 [132] ebenfalls als Quellen, da hier spezifiziert ist, was von Seiten des Gesetzgebers zur Informierung und Aufklärung in Hinblick auf den Datenschutz gefordert ist. Für die in Abbildung 7.4 abgebildeten High-Level Anforderungen haben wir vier Cluster gebildet und die jeweiligen Anforderungen entsprechend eingeordnet.

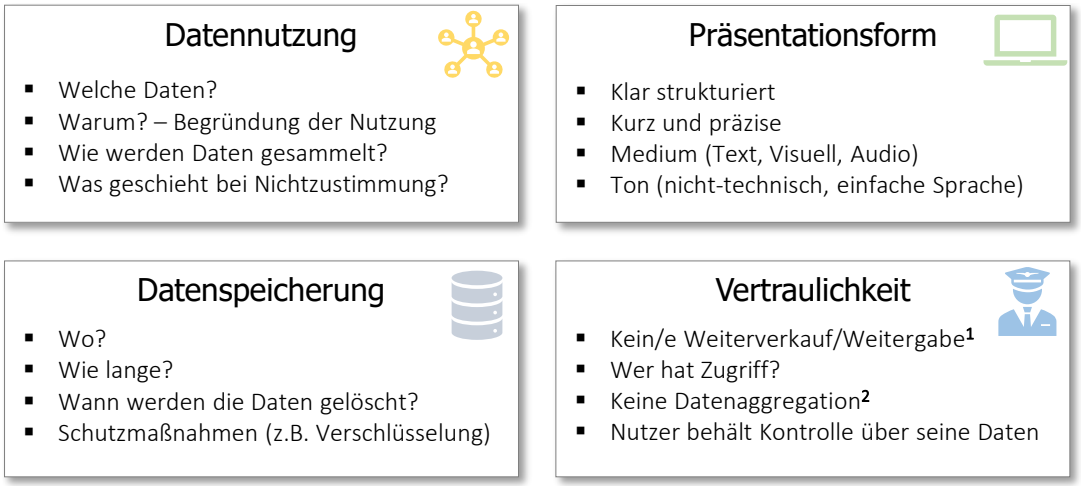


Abbildung 7.4: *High-Level* Anforderungen an Privacy Explanations

Datennutzung Hinsichtlich der Datennutzung ist es wichtig, die Benutzer darüber zu informieren und aufzuklären, welche Privatsphäreaspekte verwendet werden, warum dieses geschieht und auf welche Art und das Weise auf diese Daten zugegriffen wird [122, 164, 114, 290, 118, 399, 400]. In unseren Arbeiten [41, 390] bezeichnen wir dies als **2W1H-Prinzip**. Die zwei „W's" stammen aus den englischen Fragewörtern ***What*** und ***Why*** und meinen welche Daten gesammelt werden und warum diese benötigt werden. Das „H" rührt vom englischen Fragewort ***How*** her und bezeichnet das *Wie* die Daten gesammelt werden. Diesen Begriff verwende ich nachfolgend ebenfalls, da es sowohl den Lesefluss als auch das Verständnis verbessert. Des Weiteren reicht es nicht aus, nur über mögliche Datenpraktiken zu informieren, sondern es ist für Benutzer ebenfalls von Bedeutung, was passiert, wenn der Datenverarbeitung nicht zugestimmt wird [401, 402]. Das heißt, es muss zusätzlich erklärt werden, ob mögliche Funktionseinschränkungen vorliegen können, wenn private Daten nicht zur Verarbeitung vom Benutzer freigegeben werden.

Datenspeicherung Neben der Nutzung und Verwendung von privaten Daten ist es ebenso wichtig die Benutzer darüber aufzuklären, wie ihre Daten gespeichert werden und wird von der DSGVO auch explizit verlangt [122]. Hierzu zählt neben dem „wo“, also das Land, in dem die Daten persistent gespeichert werden, auch wie lange, das geschehen soll inklusive der Information, ob und wann die persönlichen Daten gelöscht werden, wie auch das Ergreifen von Schutz- bzw. Vorsichtsmaßnahmen (z.B.: Datenverschlüsselung) [41, 287, 189, 289, 293, 298, 110].

Vertraulichkeit Vertraulichkeit und Vertrauenswürdigkeit in Zusammenhang mit privaten Daten sind weitere essentielle Punkte, die in Anforderungen an Privacy Explanations berücksichtigt werden sollten. Wickramasinghe und Reinhardt merken an, dass es wichtig ist, Endbenutzern im „gesamten Prozess der Datenerhebung, -speicherung und -weitergabe“ die Kontrolle zu geben, um sie somit zu befähigen, kontextbezogene Entscheidungen bezüglich der Privatsphäre zu treffen [189]. Kontrolle über die eigenen Daten fördert das Vertrauen auf der Seite der Endbenutzer, was wiederum das Vertrauen in den Service Provider stärkt und ebenfalls dessen Vertrauenswürdigkeit [167, 403, 404]. Daher muss eindeutig geklärt sein, wer Zugriff hat, ob diese mit Dritten geteilt und ob die Daten möglicherweise verkauft werden [274, 114, 287, 290, 399, 110]. Des weiteren muss klar ersichtlich gemacht werden, ob die Daten Rückschlüsse auf Nutzer zulassen und ob diese ggf. mit weiteren Daten aggregiert werden, so Veys et al. [297].

Präsentationsform Eines des Hauptprobleme mit DSEs ist, dass diese zu lang, oft unverständlich und bewusst vage formuliert sowie gar nicht oder nur unzureichend strukturiert sind. Für Privacy Explanations ist es daher wichtig, dass sie gut strukturiert und verständlich formuliert (präzise und kurz) sind. Der Ton sollte nicht-technisch sein und einfache sprachliche Formulierungen verwenden [4, 42, 49, 214, 279, 114, 287, 290, 405].

7.2.3.1 Typen von Erklärungen

Miller [155] untersuchte die Erklärbarkeitsliteratur auf den Gebieten der Philosophie und Psychologie und stellte dabei fest, dass die verschiedenen Arten von Erklärungen jeweils ihre eigenen Vor- und Nachteile mit sich bringen. Diese Vor- und Nachteile müssen bei der Entwicklung von erklärbaren System abgewogen werden, um für den jeweiligen Anwendungsfall zu bestimmen, welche Art von Erklärung am besten geeignet ist, um die zuvor definierten Ziele zu erreichen. Erklärungen werden gemeinhin als Antworten auf „*W*“-Fragen (z.B. *Warum*, *Was*, *Wann*, aber auch *Wie*) beschrieben [406, 407]. Es existieren viele verschiedene Arten von

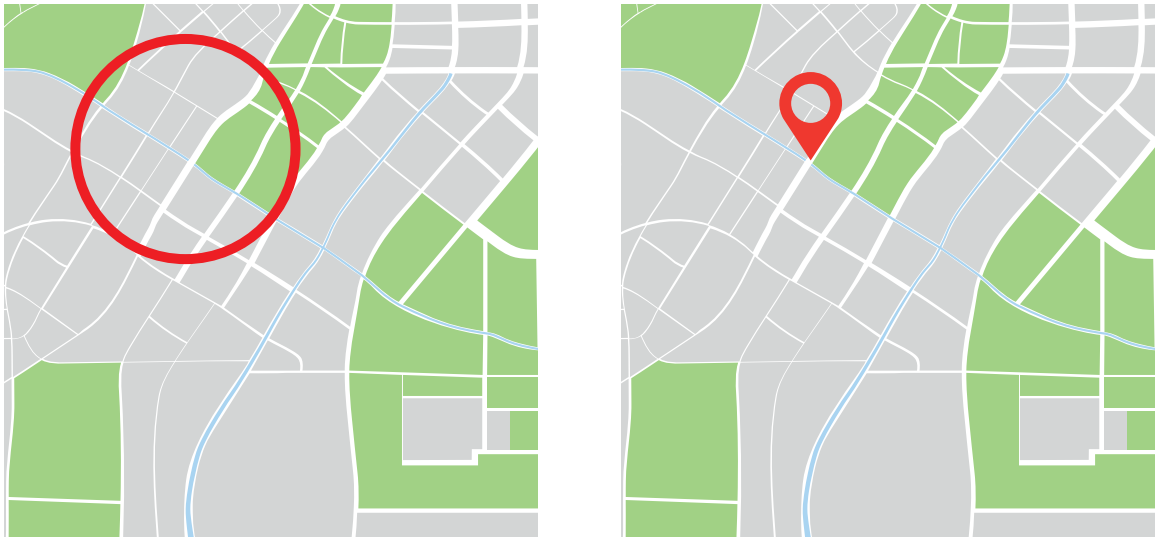
Erklärungen. Jede bringt ihre eigenen Vor- und Nachteile mit sich. Eine umfassende Übersicht über die Erklärungsarten würde im Rahmen meiner Dissertation zu weit führen, daher beschränke ich mich auf einen für Privacy Explanations wichtigen Überblick.

Eine der Herausforderungen im Umgang mit Erklärbarkeit ist, das verschiedene Nutzer unterschiedliche Bedürfnisse in Bezug auf Erklärungen haben [218, 240, 258, 260, 408, 409, 410]. Um dieser Herausforderung zu begegnen, schlagen mehrere Arbeiten aus der Erklärbarkeitsforschung personalisierte Erklärungen [240, 258, 260, 409] oder sogenannte *Layered Explanations* [41, 390, 408, 411] vor. Durch die Personalisierung bzw. der Layered Explanations können die in der Erklärung transportierten Informationen speziell auf die Bedürfnisse des einzelnen Endbenutzers zugeschnitten werden. Bei der Gestaltung personalisierter Erklärungen kann man jedoch nicht nur den Inhalt der Erklärung an den Adressaten anpassen, sondern auch die Art und Weise, wie die Informationen dem Adressaten präsentiert werden. Im Folgenden gehe ich auf zwei spezielle Arten von Erklärungen innerhalb der allgemeinen Erklärbarkeitsforschung ein, die für Privacy Explanations von großer Relevanz sind und sich von den normalen textuellen Erklärungen abheben. Anschließend stelle ich drei weitere Typen von Erklärungen vor, die im Kontext der Privacy Explanations Anwendung finden. Diese basieren im Wesentlichen auf den normalen textuellen Erklärungen im Bereich der Erklärbarkeit. Mit dem Begriff der *normalen* textuellen Erklärung fasse ich verschiedene in der Erklärbarkeitsforschung differenzierte Typen von Erklärungen zusammen, wie Causal Explanations [412, 413], Counterfactual Explanations [414] und auch Mechanistic Explanations [415]. Mechanistische Erklärungen gehören gemeinhin zu den Causal Explanations und erklären ein Phänomen, indem der Mechanismus erklärt wird, der für die Entstehung des Phänomens verantwortlich ist. Eine mechanistische Erklärung eines blutpumpenden Herzens würde beispielsweise die Teile des Herzens wie Kammern, Klappen etc. sowie die von diesen Teilen ausgeführten Vorgänge (z.B. Kontraktion und Entspannung der Kammern) und deren Organisation (das Blut fließt von den Vorhöfen zu den Klappen…) einschließen [416]. Die hier von mir vorgenommene Zusammenfassung ist aus meiner Sicht gerechtfertigt, da es beim Erklären in Bezug auf Privatsphäreaspekte primär um die Aufklärung *das* ein Privatsphäreaspekt genutzt wird, *wie* dieser genutzt und in *welchem Umfang* dies geschieht, also im wesentlichen eine Zusammenfassung von kausaler und ggf. mechanistischer Erklärung.

Contrastive Explanations Im XAI-Bereich werden kontrastive Erklärungen als Alternative zu z.B kausalen Erklärungen [214, 218, 155, 417, 418] eingesetzt. Während eine kausale Erklärung angibt, warum etwas passiert, gibt eine kontrastive Erklärung stattdessen an, warum etwas anstelle von etwas anderem passiert ist. Sie geben also einen Vergleichspunkt an, der das Verständnis unterstützen soll. In Anlehnung an [5] könnte eine kausale Erklärung beispielsweise erklären, warum der Abendhimmel rot ist. Die kontrastive Erklärung hingegen würde

erklären, warum der Abendhimmel rot und nicht blau ist. Bezugnehmend auf Miller [155] würden rein kausale Erklärungen von intelligenten Software-Systemen den Empfänger kognitiv überfordern, bedingt durch die Komplexität dieser Systeme. Bietet man dem Empfänger hier einen Vergleichspunkt in Form einer kontrastiven Erklärung an, kann die kognitive Belastung gesenkt werden, denn im Wesentlichen werden bei kontrastiven Erklärungen die falschen Erwartungen der Endnutzer mit der Realität der Situation verglichen. Diesen Mechanismus verfolgen wir ebenfalls in Bezug auf Privacy Explanations und dem adaptierten Einsatz von kontrastiven Erklärungen. Wir setzen diesen Typ von Erklärungen ein, um den Endnutzern mitzuteilen, auf welche Weise ihre Daten **nicht** verwendet werden, unabhängig davon, ob sie der Datenverarbeitung zustimmen oder nicht. Indem wir ihnen diese Informationen zur Verfügung stellen, wollen wir ihr Verständnis dafür verbessern, was tatsächlich mit ihren Daten geschieht und den Befürchtungen eines möglichen Datenmissbrauchs vorbeugen.

Example-based Explanations Eine andere Form der Informationsübermittlung an einen Adressaten ist die Bereitstellung von Example-Based Explanations. Im Kontext von XAI haben sich derartige Erklärungen als praktikable Alternative zu normalen textuellen Erklärungen erwiesen [214, 167, 419].



(a) Ungefähre Ermittlung des Standorts **(b)** Präzise Ermittlung des Standorts

Abbildung 7.5: Example-Based Explanation als Beispiel für den Standort

Example-based Explanations können in unterschiedlichen Erscheinungsformen auftreten, abhängig davon, was erklärt werden soll. Adadi und Berrada [214] haben zwei Arten identifiziert, *Prototypen und Kritikpunkte* (englisch: *prototypes and criticism*) und *kontrafaktische Erklärungen* (englisch: *counterfactual explanations*). Prototypen sind Beispiele aus einem Datensatz. Sie sollen zeigen wie der gesamte Datensatz aussehen könnte. Im Gegensatz dazu sind

Kritikpunkte Beispiele aus demselben Datensatz, die durch die zuvor gewählten Prototypen nicht gut repräsentiert sind. Im Zusammenspiel können Prototypen und die dazugehörigen Kritikpunkte einen guten Eindruck eines Datensatzes vermitteln, ohne diesen dabei zu sehr zu verallgemeinern, so Adadi und Berrada [214].

Kontrafaktische Erklärungen kommen in automatisierten Entscheidungssystemen zum Einsatz. Sie beschreiben, was getan werden muss, um die Entscheidung des Systems in eine andere Richtung zu lenken bzw. zu beeinflussen. Ein auf [214] basierendes Beispiel soll das einmal etwas verdeutlichen. Wird einem Bankkunden ein Bankkredit verweigert, möchte er vielleicht nicht nur wissen, warum er abgelehnt wurde, sondern auch was er dagegen tun kann. Die kontrafaktische Erklärung würde ihm nun diese Informationen liefern, indem sie ein Beispiel dafür liefert, was der Bankkunde tun müsste, um den angestrebten Kredit bewilligt zu bekommen. Laut Wachter et al. [420] kann es mehrere kontrafaktische Erklärungen gleichzeitig geben und ihre spezifische Relevanz von Fall zu Fall variieren.

Bei der Gestaltung von Erklärungen für Endnutzer können Example-based Explanations – insbesondere visuelle Darstellungen – ein wirksames Instrument sein. Die visuellen Beispiele müssen allerdings unter Berücksichtigung der Adressaten (Benutzer) ausgewählt werden und ein sinnvolles Maß an Komplexität aufweisen, damit sie informativ, aber dennoch für den Empfänger verständlich sind. Im Kontext der Privacy Explanations lassen sich Example-based Explanations sinnvoll einsetzen, um zum Beispiel technische oder auch abstrakte Sachverhalte leicht verständlich zu transportieren. Abbildung 7.5 veranschaulicht eine Example-based Explanation für eine Standortabfrage. In Abbildung 7.5a wird dem Benutzer anhand eines leicht verständlichen Schaubildes gezeigt, was eine ungefähre Ermittlung seines Standorts bedeuten würde, es wird also nur ein ungefährer Radius ermittelt, in dem er sich aufhält. Bei der präzisen Lokalisierung hingegen wird der exakte Standort vom System ermittelt (siehe Abbildung 7.5b). So ist der Benutzer, auch ohne technisches Hintergrundwissen schnell in der Lage, den Unterschied zu verstehen und diese Information in seine Entscheidungsfindung einzubeziehen.

Allgemeine textuelle Erklärungen Die nachfolgenden drei Arten von Erklärungen sind keine in der Erklärbarkeitsliteratur vorkommenden Begriffe, sondern wurden von uns geprägt [390]. Mit der **Baseline Explanation** werden dem Benutzer die ersten, grundlegenden Informationen zur Verwendung seiner Privatsphäreaspekte erklärt, gemäß dem 2W1H-Prinzip. Darüber hinaus wird erklärt, welche Konsequenzen die Nichtzustimmung der Verwendung der eigenen privaten Daten hat. Die **Detailed Explanation** vermittelt, wie es die DSGVO verlangt, feingranularere Details zur Datennutzung und die **Third Parties Explanation** gibt Aufschluss über eventuelle Nutzung der Daten von Seiten Dritter (siehe Abschnitt 7.2.3.2).

7.2.3.2 Layered-Ansatz

Um möglichst vielen Nutzern mit unterschiedlichen Haltungen zum Datenschutz gerecht zu werden, ohne aber deren Präferenzen vorab in einem System speichern zu müssen und dadurch den Aufwand für die Nutzer so gering wie möglich zu halten, habe ich das Konzept der Layered Erklärungen für Privacy Explanations adaptiert. Ein weiterer großer Vorteil der Layered Explanations ist, dass sich die hohe Informationsdichte und die damit verbundene Komplexität der Informationen, die dem Benutzer zur umfassenden Aufklärung durch die Privacy Explanations transportiert werden müssen, durch die verschiedenen Ebenen (englisch: *layer*) granular aufbereitet werden können [411, 421]. Die verschiedenen Ebenen des Informationserhalts und der -verarbeitung tragen zu einem besseren Verständnis bei [421] und adressieren, ähnlich zu personalisierten Erklärungen, ebenfalls unterschiedliche Anforderungen verschiedener Nutzer [408, 411]. Das heißt, dass das von mir vorgestellte Konzept der Privacy Explanations durch den Layered-Ansatz in Verbindung mit den verschiedenen Typen von Erklärungen in der Lage ist, alle Informationen, beginnend bei der Datennutzung bis hin zur feingranulareren Informationen über Datennutzung durch Drittanbieter, zugeschnitten auf den persönlichen Informationsbedarf eines Benutzers, was seine eigene Privatsphäre betrifft, vermitteln zu können.

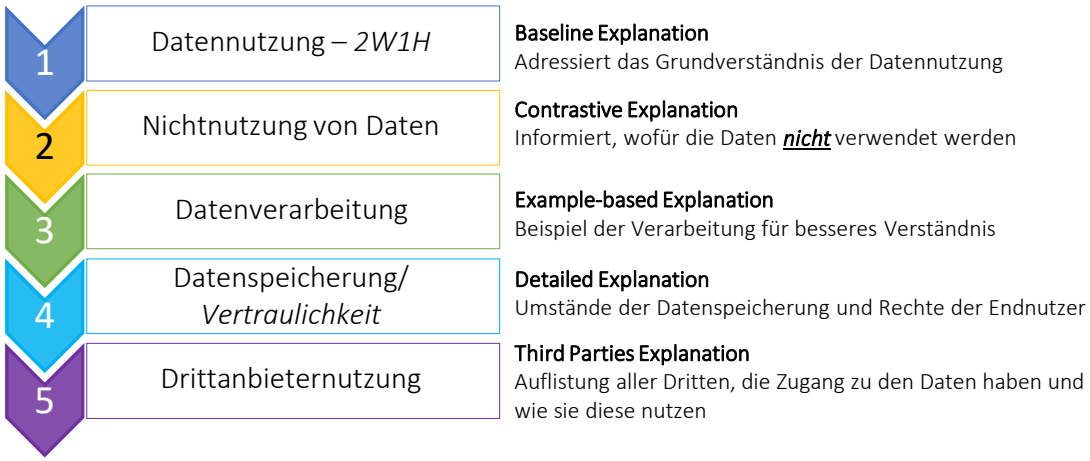


Abbildung 7.6: Überblick des Layered-Ansatzes

Abbildung 7.6 veranschaulicht den vorgestellten Layered-Ansatz der Privacy Explanations. Auf der Grundlage unserer definierten Anforderungen (Abschnitt 7.2.3) an Privacy Explanations führen wir fünf Ebenen von Erklärungen ein. Jede Ebene deckt verschiedene Datenschutzaspekte ab, und insgesamt enthalten sie alle geforderten Hinweise zum Datenschutz.

Die ersten beiden Ebenen der Privacy Explanation konzentrieren sich auf das Anforderungs-Cluster der *Datennutzung*, wobei normale textuelle Erklärungen und kontrastive

Erklärungen verwendet werden. In der ersten Ebene wird den Endnutzern gemäß dem *2W1H-Prinzip* erklärt, welche Daten verwendet werden, warum sie verwendet werden und was im Falle einer Nichtzustimmung geschieht. Die **Baseline Explanation** ist entscheidend dafür, dass die Endnutzer verstehen wie ihre Daten verwendet werden, und dass sie entscheiden können, ob sie dieser Verarbeitung zustimmen oder nicht. Die zweite Ebene der Privacy Explanation ist eine **Contrastive Explanation**, die, wie der Name schon sagt, im Gegensatz zur Baseline Explanation steht. Auf der kontrastiven Erklärungsebene wird dem Nutzer mitgeteilt, auf welche Weise dessen privaten Daten ***nicht*** verwendet werden, sofern dieser der Verarbeitung seiner Daten zustimmen sollte.

Eine bewährte *Präsentationsform*, die den Endnutzern zusätzlichen Kontext bieten kann, sind **Example-based Explanations**. Adadi und Berrada [214] stellten fest, dass „unter den agnostischen Methoden ist die Visualisierung die am meisten auf den Menschen ausgerichtete Technik". In der dritten Ebene der Privacy Explanations entscheiden wir uns daher für Example-based Explanations. Unklare Erklärungen darüber, wie Daten verarbeitet werden können zu Entscheidungen führen, die nicht den tatsächlichen Präferenzen und Haltungen der Endnutzer entsprechen. Daher sollte die Bereitstellung eines visuellen Beispiels ein effizienter Weg sein, dieses Problem zu lösen. Darüber hinaus kann der Software Engineer die Auswahl von Beispielen, die zum Software-System und seinen typischen Anwendungsfällen passen (kontextbezogen), zusätzlichen Kontext für den Endnutzer schaffen.

Im vierten und fünften Layer werden feingranularere Informationen über zusätzliche Datenschutzaspekte für Endbenutzer bereitgestellt, die ein höheres Privatsphärebedürfnis haben und somit weitere, umfassendere Informationen einfordern. Im Wesentlichen decken diese Schichten Details zur *Datenspeicherung* und *Vertraulichkeit* ab, wie sie in den Anforderungen an Privacy Explanations beschrieben sind. Der vierte Layer erklärt die Umstände der Datenspeicherung und die Rechte der Endnutzer in der **Detailed Explanation**. Die fünfte und letzte Ebene besteht aus der **Third Parties Explanation**, die Informationen zu allen Drittparteien aufweist, die Zugang zu den Daten erhalten würden sowie kurz angibt, wie sie diese verarbeiten könnten [122, 274].

7.2.4 Umsetzung in einem ersten Prototyp

Um die praktische Anwendbarkeit unseres durch Literatur und Umfragedaten gestützten theoretischen Konzepts der Privacy Explanations zu evaluieren, haben wir dieses in einem interaktiven Prototypen umgesetzt. Der Fokus bei der Gestaltung des Prototypen lag auf der Umsetzung des theoretischen Konzepts, also Arten der Erklärungen, deren Lesbarkeit sowie deren Verständlichkeit. Daher haben wir das Design bewusst sehr einfach gehalten und nicht

in die Umgebung einer Anwendung eingebettet. Des Weiteren haben wir zwei unterschiedliche Sprachkonzepte verwendet, die direkte Ansprache, die den Nutzer direkt anspricht (*deine Daten* oder *deine Rechte*) sowie eine formalere Ansprache, die sprachliche Konstrukte wie *Nutzerdaten* oder *Nutzerrechte* verwendet. Innerhalb des Prototypen haben wir zu fünf Privatsphäreaspekten (Ungefährer Standort, Genauer Standort, Geburtsdatum, Bilddateien und Browser-Verlauf) jeweils unsere fünf verschiedenen Erklärungstypen verwendet, mit Ausnahme des Geburtsdatums. Hier haben wir keine Example-based Explanation gegeben, da jedem im Grunde intuitiv klar ist, was ein Geburtsdatum ist. Anders schaut es beim Standort aus. Laut Fu und Lindqvist [400] haben Endbenutzer im Allgemeinen ein gutes Verständnis was mit einem *präzisen* (genauen) Standort gemeint ist. Hingegen ist das Verständnis für einen *ungefähren* Standort häufig nicht gegeben.

7.2.4.1 Architektur und Implementation des Prototypen

Zum Evaluieren der einzelnen Privacy Explanations wurde bei der Architektur das Konzept der Kompartmentalisierung angewandt. Jeder Privatsphäreaspekt war einzeln über Tabs abrufbar. Innerhalb der jeweiligen Privatsphäreaspekte konnte der Nutzer durch die verschiedenen Typen von Erklärungen navigieren. Abbildung 7.7 veranschaulicht hierzu die Navigationsrouten.

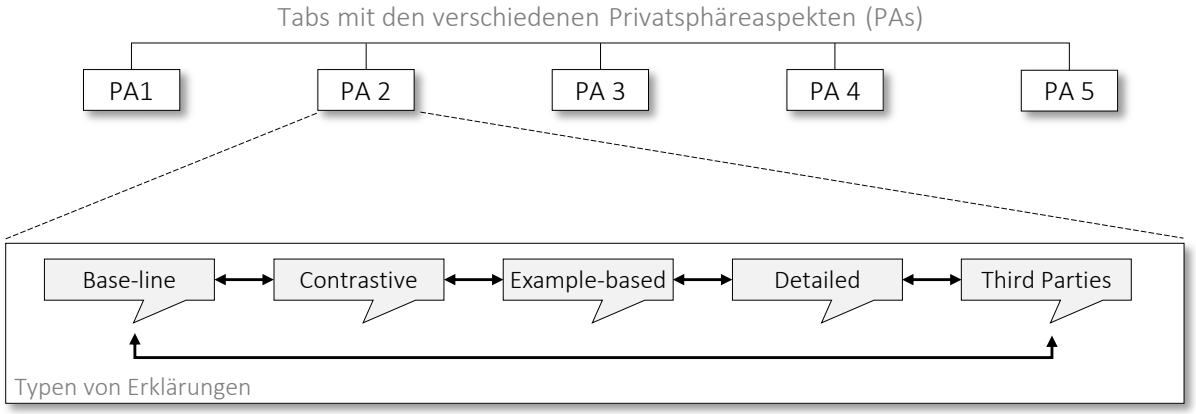


Abbildung 7.7: Navigationsrouten für den Layered-Ansatz innerhalb des Prototyps in Anlehnung an [5]

Abbildung 7.8 stellt die Oberfläche des Prototyps dar, die mit der Software Axure RP[3] umgesetzt wurde. ① zeigt die fünf Tabs mit den jeweiligen Privatsphäreaspekten. ② bildet die Base-line Explanation mit der Datennutzung und dem Hinweis ab, was bei Nichtzustimmung geschieht ab. Durch den Toggle Button bei ③ kann der Nutzer seine individuelle Zustimmung oder Nichtzustimmung für jeden Privatsphäreaspekt einzeln geben. Das ermöglicht es den Nutzern die tatsächliche Kontrolle über ihre persönlichen Daten auszuüben, anstatt

[3] `https://www.axure.com`

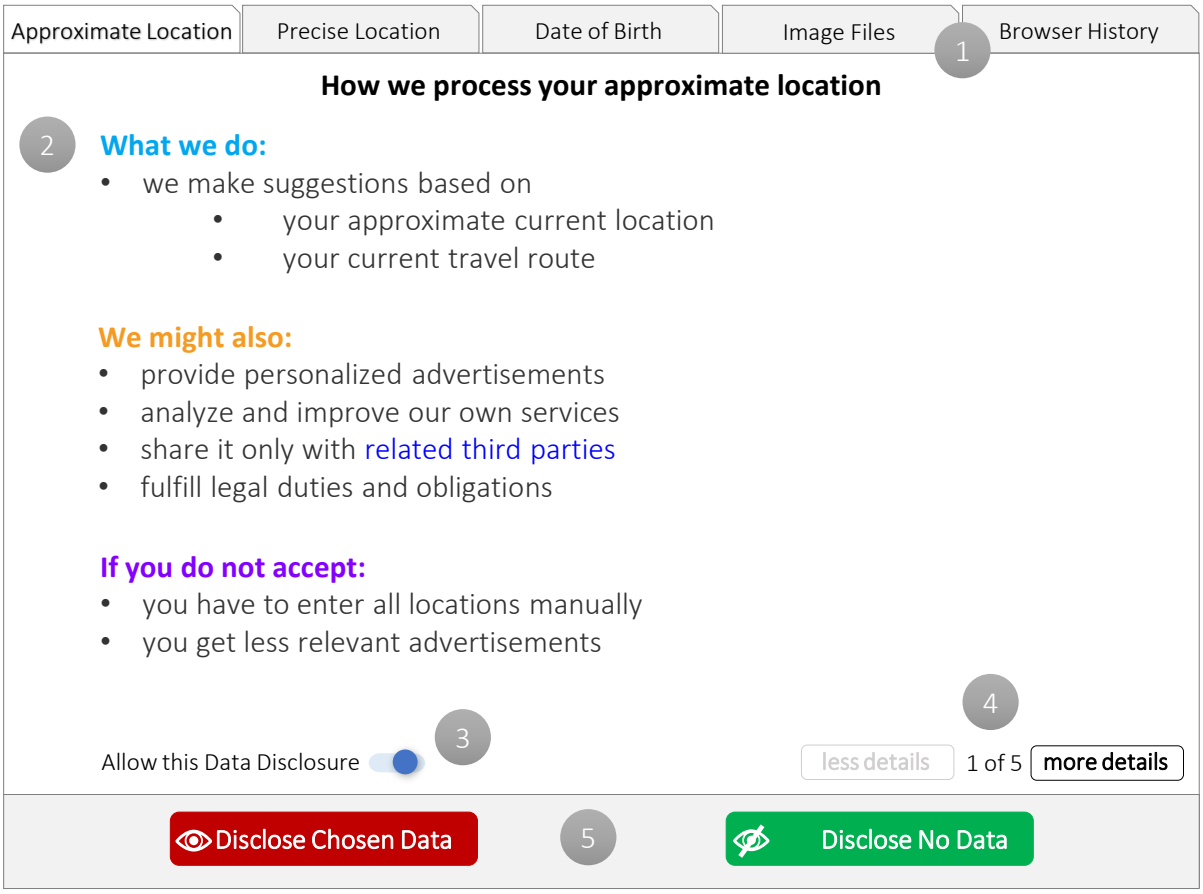


Abbildung 7.8: Navigationsrouten innerhalb des Prototypen in Anlehnung an [5]

sie zu zwingen, sich zwischen der eigenen Privatsphäre und Service-Angebot entscheiden zu müssen [116]. Mit den „less details“ und „more details“ Buttons (④) kann der Benutzer durch die in Abbildung 7.7 erwähnten Routen navigieren, um die verschiedenen Ebenen der Privacy Explanations einzusehen. ⑤ zeigt die Buttons mit denen der Benutzer seine Entscheidungen bezüglich der Offenlegung seiner Privatsphäreaspekte treffen kann. Hier kann er entweder der nur von ihm explizit gewählten Offenlegung zustimmen (Toggle Button ③) oder sämtlicher Offenlegung seiner Daten widersprechen. Letzteres erfordert keinerlei Interaktion mit dem Prototyp, insbesondere keiner Deaktivierung der einzelnen Toggle Buttons. Dadurch setzen wir in unserem Prototyp konsequent das *Opt-In*-Verfahren um. Opt-In bedeutet, dass der Datenbesitzer explizit zustimmen muss (englisch: *explicit consent*), bevor seine Daten erfasst und verarbeitet werden. „Opt-in ist fairer, und die Nutzung des Dienstes sollte nicht von der Zustimmung zur Datenerfassung abhängig gemacht werden“, so Bruce Schneier [31, S. 198]. Denn im Gegensatz dazu werden beim *Opt-Out*-Verfahren Daten erfasst und gesammelt ohne evtl. Wissen und ohne vorherige Zustimmung des Datenbesitzers. Dieser kann im Nachhinein der Datenerfassung widersprechen. Das ist aber in der Praxis häufig nicht ganz einfach, denn

entsprechende Optionen dazu sind meistens schwer zu finden, in dem Wissen, dass die Nutzer sich nicht darum extra bemühen würden [31].

In Abschnitt E.2 sind weitere Screenshots des Prototyps abgebildet samt der verschiedenen Typen von Erklärungen, sowie weitere Erläuterungen bezüglich Wahl von Icons und der Anwendung umgekehrter Dark Patterns im Prototypen.

7.2.5 Evaluation

Nachdem wir nun, gestützt durch Literatur und Umfragedaten, ein umfassendes Konzept für Privacy Explanations entwickeln konnten, sollte dies in einer Benutzerstudie getestet werden. Der Fokus der Studie lag auf der Evaluation des von uns entwickelten Konzepts und nicht auf der Usability des Prototypen. Daher wurde im Rahmen dieser Studie die subjektiven Erfahrungen und Wahrnehmungen der Teilnehmer im Umgang mit dem Prototyp untersucht.

7.2.5.1 Design der Studie

Die Studie bestand aus einer Kombination von semistrukturierten Interviews und dem praktischen Arbeiten mit dem Prototypen (Abschnitt 7.2.4) durch die Teilnehmer. Zu Beginn wurden Daten zum demografischen Hintergrund der Teilnehmer abgefragt sowie zu deren Bewusstsein für die eigene Privatsphäre. Anschließend sollten die Probanden durch den Prototypen navigieren. Um sich besser in die Situation hineinversetzen zu können, wurde den Teilnehmern dafür zuvor ein hypothetisches Szenario an die Hand gegeben. Das verwendete Szenario ist in Abschnitt E.3 zu finden. Anschließend beantworteten die Teilnehmer Fragen bezüglich des Prototypen sowie Fragen (offene und geschlossene Fragen) zum Einfluss der Studie auf ihre Privacy Awareness. Die Studie wurde im Vorfeld mit zwei Doktoranden des FG Software Engineering pilotiert. Da wir die verschieden sprachlichen Erklärungen testen wollten, haben wir den Teilnehmern randomisiert entweder jeweils die Erklärungen mit direkter Ansprache gezeigt oder die, mit der formaleren Ansprache. Aufgrund der andauernden Pandemie war es nicht möglich, die Studie in Präsenz durchzuführen, so dass wir diese Online durchführten.

7.2.5.2 Durchführung der Studie

Bei der Durchführung der Online-Studie haben wir AnyDesk verwendet, um den Teilnehmern Zugang zum Prototyp zu gewährleisten. Als Voice Chat-Anwendung kamen BigBlueButton, Skype, Discord[4] oder Facetime[5] zum Einsatz sowie OBS Studio zum Aufzeichnen von Video- und Audio-Daten. Die Teilnehmer wurden sowohl bei der Interaktion mit dem Prototyp, als

[4] https://discord.com

[5] https://apps.apple.com/de/app/facetime/id1110145091

auch beim Bearbeiten der Fragebögen gebeten, die *Think Aloud*-Methode zu verwenden. Zudem fragte der Interviewer bei unklaren oder unerwarteten Äußerungen nach, so dass die Probanden hier ihre Entscheidungen oder Äußerungen genauer begründen mussten.

7.2.5.3 Datenerhebung und Analyse

Wir haben die Studie zwischen Januar und Februar 2022 durchgeführt. Die Studie wurde mit insgesamt 61 Probanden durchgeführt. Davon wurden 29 Teilnehmern die Erklärungen mit direkter Ansprache gezeigt und die übrigen 32 Teilnehmer erhielten Erklärungen in indirekter Ansprache. Die Teilnehmerakquise fand über akademische Kontakte und Anfragen an Studierende der Leibniz Universität Hannover statt. Die Analyse der erhobenen Daten erfolgte sowohl quantitativ als auch qualitativ. Für die qualitative Analyse wurde neben den Daten aus dem Fragebogen auch das entstandene Videomaterial vollständig gesichtet und kodiert. Die Kodierung bestand aus insgesamt zwei Zyklen. Beginnend mit *In Vivo*-Kodierung nach Charmaz [231] und anschließendem Pattern Coding [232]. Der Kodierungsprozess wurde unabhängig voneinander (von Jakob Droste und mir) durchgeführt. Bei Unstimmigkeiten diskutierten die Autoren diese, bis ein gemeinsamer Konsens erreicht wurde. Für die quantitative Analyse haben wir statistische Signifikanztests durchgeführt. Die zusammengefassten Nullhypothesen dazu sind in Tabelle 7.1 zusammengefasst.

Tabelle 7.1: Nullhypothesen für die quantitative Datenanalyse

ID	Hypothesen
$H1_0$	Es gibt keinen signifikanten Unterschied zwischen direkter und indirekter Ansprache, wenn es um die Vorliebe von Privacy Explanations geht.
$H2_0$	Es gibt keinen signifikanten Unterschied zwischen direkter und indirekter Ansprache, wenn es um die wahrgenommene Glaubwürdigkeit von Privacy Explanations geht.
$H3_{0,i}$	Es besteht kein signifikanter Unterschied zwischen der Base-line Explanation und den anderen Erklärungstypen hinsichtlich ihrer Relevanz.
$H4_{0,i}$	Es gibt keinen signifikanten Unterschied zwischen der Base-line Explanation und den anderen Erklärungstypen hinsichtlich ihrer Verständlichkeit.
$H5_{0,i}$	Es gibt keinen signifikanten Unterschied zwischen direkter und indirekter Ansprache, wenn es um die wahrgenommenen Faktoren F_i. $F_{security}$: Empfundene Sicherheit der eigenen Daten $F_{privacy}$: Empfundene Wahrung der Privatsphäre $F_{trustworthiness}$: Glaubwürdigkeit der Erklärungen $F_{readiness}$: Bereitschaft die Software einzusetzen

7.2.5.4 Ergebnisse der Studie

Demographische Daten und Privacy Awareness Das Alter der Teilnehmer reicht von 19 bis 60 Jahren (M=25, SD=9,93). Sofern nicht anders angegeben, stehen die Abkürzungen M für das arithmetische Mittel und SD für die Standardabweichung. Alle Teilnehmer nutzen das Internet täglich. 46% gaben sogar an, mehr als 5 Stunden täglich das Internet zu nutzen. 87% der Teilnehmer gaben an, aktiv an den sozialen Netzwerken zu partizipieren. Zu Beginn der Studie wurden die Teilnehmer bezüglich Ihrer Privacy Awareness befragt. 82% ist der Schutz ihrer persönlichen Daten wichtig (34,4% sehr wichtig). Lediglich einem Teilnehmer ist das nicht so wichtig und den restlichen 16,3% ist es einigermaßen wichtig. Tabelle 7.2 zeigt das Ergebnis auf die Frage „Wie wichtig ist es ihnen zu wissen, ***wer*** ihre persönlichen Daten verarbeitet?“ und Tabelle 7.3 die Ergebnisse zur Frage „“Wie wichtig ist es ihnen zu wissen, ***welche*** ihrer persönlichen Daten verarbeitet werden?.

Tabelle 7.2: Wie wichtig ist es ihnen zu wissen, ***wer*** ihre persönlichen Daten verarbeitet?
Legende: 5 - Sehr wichtig, 4 - Wichtig, 3 - Einigermaßen wichtig, 2 - Nicht so wichtig, 1 - Unwichtig

Kategorie	**5**	**4**	**3**	**2**	**1**
Smartphone Apps	36,1%	31,1%	23,0%	8,2%	1,6%
Software Arbeit	23,0%	34,4%	19,7%	19,7%	3,3
Software Privat	27,9%	42,6%	16,4%	9,8%	3,3%
Staatliche Behörden	26,2%	32,8%	13,1%	24,6	3,3%
Websites	32.8%	37,7%	19,7%	8,2%	1,6%

Tabelle 7.3: Wie wichtig ist es ihnen zu wissen, ***welche*** ihrer persönlichen Daten verarbeitet werden?
Legende: 5 - Sehr wichtig, 4 - Wichtig, 3 - Einigermaßen wichtig, 2 - Nicht so wichtig, 1 - Unwichtig

Kategorie	**5**	**4**	**3**	**2**	**1**
Adressdaten	72,1%	21,3%	4,9%	1,6%	–
Bankdaten	91,8%	6,6%	1,6%	–	–
Fotos & Bilder	60,7%	26,2%	13,1%	–	–
Hobbies/Interessen	6,6%	24,6%	41,0%	23,9%	4,9%
Medizinische Daten	80,3%	16,4%	3,3%	–	–
Standort-Daten	42,6%	36,1%	14,8%	4,9%	1,6%
Suchverhalten	24,6%	27,9%	23,0%	21,3%	3,3%
Surfverhalten	24,6%	32,8%	24,6%	14,8%	3,3%

Die Ergebnisse deuten darauf hin, dass den Teilnehmern ihre Privatsphäre wichtig ist und sie ein Bewusstsein dafür haben. Das wird bekräftigt dadurch, dass sich 89% der Teilnehmer angaben zu wissen, dass ihre persönlichen Daten mit Drittanbietern geteilt werden (in Bezug auf Wikipedia, Amazon und Instagram) und die übrigen 11% äußerten, dass sie es nicht explizit

wussten, aber davon ausgegangen sind, dass dies der Fall ist. Zudem geht aus den Daten hervor, dass die Studienteilnehmer verschiedenen Privatsphäreaspekten einen unterschiedlichen Wert in Bezug auf dessen Schutz beimessen und im Allgemeinen über dessen Verwendung informiert werden möchten.

Anzahl, Gliederung und Länge der Privacy Explanations Nachdem die Teilnehmer mit dem Prototyp interagiert haben, hatten sie einige Fragen dazu zu beantworten. Zunächst sollten die Teilnehmer Auskunft geben wie gut ihnen die Erklärungen generell gefallen haben in Bezug auf Anzahl der Layered-Erklärungen sowie dessen Gliederung. Abbildung 7.9 stellt das Ergebnis dar. Die Ergebnisse für die Anzahl der Datenschutzerklärungen (M=4,21, SD=1,01003) deuten darauf hin, dass sich die Teilnehmer von der Anzahl der Erklärungen nicht überfordert fühlten. Die Reihenfolge der Privacy Explanations (M=4,23, SD=0,89431) wurde insgesamt als gut empfunden.

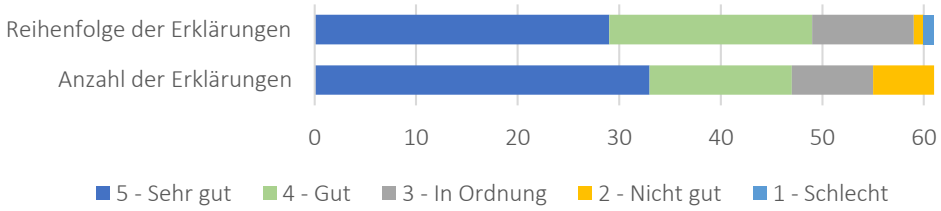


Abbildung 7.9: Ergebnis Anzahl und Reihenfolge der Datenschutzerklärungen

Die Teilnehmer wurden ebenfalls nach ihrer Meinung gefragt, was die Länge bzw. den Umfang der verschiedenen Privacy Explanations-Typen anging. Hierfür wurden für jeden Erklärungstyp einzeln die Wahrnehmungen der Probanden abgefragt. Das Ergebnis hierzu ist in Abbildung 7.10 dargestellt.

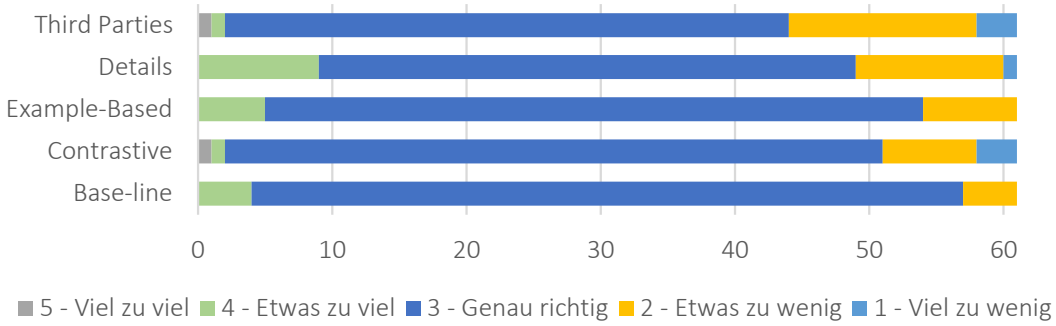


Abbildung 7.10: Umfang (Länge) der unterschiedlichen Erklärungstypen

Insgesamt zeigt Abbildung 7.10, dass der Umfang der verschiedenen Erklärungstypen als gut empfunden wurden. Sowohl Base-line als auch Example-based zeigen ein eher ausgewogenes Bild. Die Typen Contrastive und Third Parties zeigen hingegen, dass Sie als eher etwas

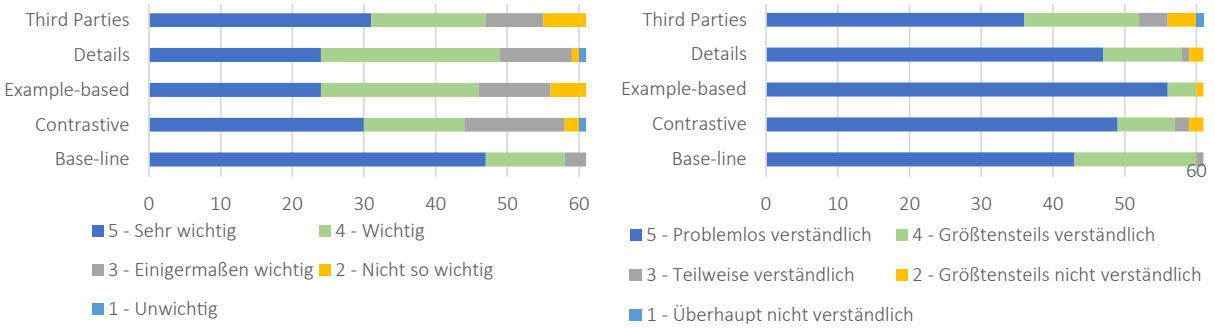
zu kurz wahrgenommen wurden. Eine Contrastive Explanation fällt im Gegensatz zu einer Base-line Explanation i.d.R. deutlich kürzer aus, da sie lediglich angibt, wofür ein spezifischer Privatsphäreaspekt nicht genutzt wird (siehe Screenshot Abbildung E.2). Dieser Effekt könnte hier eine Rolle gespielt haben, zumal dieser auch von einigen Probanden explizit erwähnt wurde. Bei den Erklärungen zu den Drittanbietern (Abbildung E.5) wurden lediglich Verweise inklusive Hyperlink auf den Anbieter samt kurzer Erläuterung, was der Grund für die Teilung der Daten ist, gegeben. Hier äußerten einige Teilnehmer später, dass sie hier gerne spezifischere Details zur Drittanbieternutzung gehabt hätten.

Unterschiede in Ansprache, Glaubwürdigkeit und bei Erklärungstypen Um festzustellen, ob es einen Unterschied zwischen den unterschiedlichen Ansprachen der Teilnehmer in den Erklärungen in Bezug auf Gefallen/Vorlieben sowie die Glaubwürdigkeit der einzelnen Erklärungen gab, sollten die Teilnehmer auf einer 5 Punkte Likert-Skala (5 - Sehr gut bis 1 - Schlecht) ihre Meinung abgeben. Wir haben jeweils mit dem Wilcoxon-Mann-Whitney-Test [422, 423] auf statistische Signifikanz getestet. Sowohl was die Vorliebe der einzelnen Teilnehmer anging als auch die Glaubwürdigkeit in Zusammenhang mit der verwendeten Ansprache, konnte keine Signifikanz nachgewiesen werden (siehe Tabelle 7.4 für Signifikanztests). Im Allgemeinen hat der Inhalt der Privacy Explanations den Teilnehmern gut gefallen (M=4,46, SD=0,66711), wie in Abbildung E.6a dargestellt ist. Die Glaubwürdigkeit wurde ebenfalls als gut von den Teilnehmern eingestuft (M=4,05, SD=1,01507), wie in Abbildung E.6b zu sehen.

Tabelle 7.4: Ergebnisse der Signifikanztests

ID	Ergebnisse	Signifikant?
$H1_0$	$z \approx 0,09387,\ p \approx 0.92828$	Nein
$H2_0$	$z \approx -0,38993,\ p \approx 0,69654$	Nein
$H3_{0,\ contrastive}$	$z \approx -3,5797,\ p \approx 0,00034$	Ja
$H3_{0,\ example-based}$	$z \approx -6,3931,\ p < 0,00001$	Ja
$H3_{0,\ details}$	$z \approx -3,9719,\ p \approx 0,00008$	Ja
$H3_{0,\ thirdparties}$	$z \approx -3,8097,\ p \approx 0,00014$	Ja
$H4_{0,\ contrastive}$	$z \approx -0,3823,\ p \approx 0,70394$	Nein
$H4_{0,\ example-based}$	$z \approx -2,334,\ p \approx 0,0198$	Ja
$H4_{0,\ details}$	$z \approx -0,0284,\ p \approx 0,97606$	Nein
$H4_{0,\ thirdparties}$	$z \approx -2,5732,\ p \approx 0,01016$	Ja
$H5_{0,\ F_{security}}$	$\chi^2 \approx 1,0603,\ p \approx 0,588521$	Nein
$H5_{0,\ F_{privacy}}$	$\chi^2 \approx 2,3247,\ p \approx 0,312743$	Nein
$H5_{0,\ F_{trustworthiness}}$	$\chi^2 \approx 1,6183,\ p \approx 0,445141$	Nein
$H5_{0,\ F_{readiness}}$	$\chi^2 \approx 2,4161,\ p \approx 0,298785$	Nein

Für die Signifikanztests der Hypothesen H3 und H4 wurde der Wilcoxon-Vorzeichen-Rang-Test [422] eingesetzt. Es wurde hier jeweils verglichen zwischen Base-line Explanation und einem der anderen vier Erklärungstypen. Bezüglich der Relevanz konnte in allen vier Fällen eine statistische Signifikanz festgestellt und somit die Nullhypothesen $H3_0, i$ verworfen werden. Die Teilnehmer waren eindeutig der Meinung, dass die Base-line Explanation für sie am wichtigsten war (M=4,72, SD=0,54716). Mit fast gleichen Mittelwerten rangierten die Contrastive (M=4,15, SD=0,98122), die Details (M=4,15, SD=0,86529) und die Third Parties Explanation (M=4,18, SD=1,00013) an zweiter Stelle. Die beispielhafte Erklärung wurde eindeutig als am wenigsten wichtig empfunden (M=3,07, SD=0,93859). Abbildung 7.11a veranschaulicht dies. Ein interessanter Fakt ist allerdings, wenn es um die Präferenzen der Teilnehmer bei Privacy Explanations geht, dass der Code „*hilfreiche Beispiele geben*“ mit 26 Nennungen der am meist genannte ist (siehe Tabelle E.1). Das schlechtere Abschneiden im Prototyp lässt sich möglicherweise darauf zurückführen, dass die Teilnehmer sich schnell und einfach vorstellen konnten, was die Nutzung eines spezifischen Privatsphäreaspekts angeht. Gestützt wird dies dadurch, dass die Example-based Explanation als der am besten verständlichste Erklärungstyp abschnitt. Sogar signifikant besser als die Base-line Explanation (Tabelle 7.4). Infolgedessen hat die Example-based Explanation durchaus ihre Berechtigung und sollte bei der Implementation von Privacy Explanations berücksichtigt werden.



(a) Relevanz der Erklärungstypen **(b)** Verständlichkeit der Erklärungstypen

Abbildung 7.11: Überblick von Relevanz und Verständlichkeit der Erklärungstypen

Beim Vergleich der Verständlichkeit konnte nur für die Example-based und Third Parties Explanation statistische Signifikanz nachgewiesen werden. Das bedeutet, dass die subjektive Verständlichkeit der Contrastive und Details Explanation nicht signifikant anders empfunden wurde, als die der Base-line Explanation. Ein qualitativer Vergleich des Verständnisses ist in Abbildung 7.11b abgebildet. Die Base-line Explanation schien die Teilnehmer beim Verständnis nicht vor große Probleme zu stellen (M=4,69, SD=0,49724). Die überwiegende Mehrheit der Teilnehmer gab an, zumindest die meisten der darin enthaltenen Informationen

verstanden zu haben. Die Example-based Explanation lag, was das Verständnis angeht, signifikant ganz vorne (M=4,89, SD=0,44715). Die Third Parties Explanation wurde als signifikant weniger verständlich wahrgenommen als die Base-line Explanation (M=4,34, SD ≈ 0,97317). Die Verständlichkeit der Contrastive (M ≈ 4,7, SD=0,68579) und der Details Explanation (M ≈ 4,69, SD=0,66631) unterschied sich nicht signifikant von der der Base-line Explanation.

Nachdem die Teilnehmer durch den Prototyp navigiert hatten, wollten wir deren Wahrnehmungen und Empfindungen in Bezug auf die Interaktion mit dem Prototyp wissen. Genauer gesagt wurden sie gefragt, ob die Privacy Explanations einen Einfluss auf die folgenden in Tabelle 7.1 genannten Faktoren F_i hatten (Security, Privacy, Trustworthiness und Readiness). Die Einflüsse auf diese Faktoren wurden ebenfalls mittels des Between-Groups-Designs getestet. Hierzu konnten die Teilnehmer einen möglichen Effekt mit *positiv*, *negativ* oder *weder noch* angeben. Da wir nominale Daten haben, setzten wir den Chi-Quadrat-Test zur Prüfung auf Signifikanz ein. Wie in Tabelle 7.4 zu sehen, konnte für keinen der Fälle statistische Signifikanz nachgewiesen werden. Somit lässt sich aus den Daten für unsere Teilnehmer schließen, dass die Art der im Prototyp verwendeten Formulierung keinen signifikanten Einfluss auf die Faktoren F_i haben. Abbildung 7.12 zeigt qualitativ die Effekte für beide Gruppen. Im Anhang in Abbildung E.7a und Abbildung E.7b sind die Effekte für jede Gruppe einzeln dargestellt.

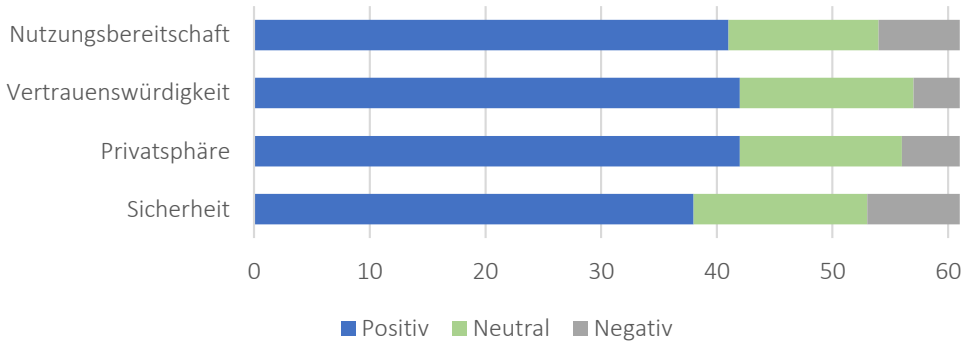


Abbildung 7.12: Überblick der Einflüsse von Privacy Explanations (beide Gruppen)

Auch wenn es keine signifikanten Unterschiede zwischen beiden Gruppen gab, deuten die Gesamtergebnisse darauf hin, dass die Mehrheit der Teilnehmer Privacy Explanations als positiv für alle untersuchten Faktoren empfanden. Auf die Nachfrage, was den Teilnehmern, die hier mit *negativ* gestimmt missfiel, gaben diesen an, dass es spezifischer Inhalt einiger Erklärungen war, der ihnen nicht gefallen hat. Zum Beispiel gefiel mehreren dieser Teilnehmer der Umfang der persönlichen Daten nicht, auf die die App des hypothetischen Szenarios zugreifen wollte. Es also nicht an den Privacy Explanations an sich lag.

Eignung der Erklärungstypen Insgesamt wurde das Konzept der Privacy Explanations von den Teilnehmern positiv bewertet. Rund 44,3% der Probanden war dieses Konzept gänzlich neu und sie konnten sich auch zu Beginn nichts unter einer Privacy Explanation vorstellen. 39% gaben an, eine vage Idee des Konzeptes zu haben, konnten dies aber nicht konkretisieren. 55,7% der Teilnehmer gaben an, durch die Privacy Explanations für den Datenschutz sensibilisiert worden zu sein, erlebten also einen Zuwachs Ihrer Privacy Awareness. 27,9% verneinten dies. Interessant ist allerdings, dass 8 von diesen 17 Teilnehmern, angab, bereits eine hohe Privacy Awareness zu haben.

Der Layered-Ansatz, der zum einen die Komplexität und Menge an Informationen zugänglicher machen soll, sowie die darin eingebetteten unterschiedlichen Typen von Erklärungen wurden von den Teilnehmern als geeignetes Mittel zur Aufklärung über Datennutzung und Datenpraktiken bei der Verwendung von Privatsphäreaspekten bewertet (siehe Abbildung 7.9 und Abbildung 7.10). Der nächste Schritt ist nun die Umsetzung der Konzeptstudie sowie deren Implementation in einen Software-Prototyp.

7.3 Technische Umsetzung und Evaluation

Auf Grundlage der gewonnenen Daten haben wir begonnen, das Konzept der kontextuellen Privacy Explanations technisch umzusetzen. Dazu wurde ein Anwendungskontext definiert, der einem in der realen Welt möglichst weit verbreiteten und bekannten Szenario entspricht. Die Wahl fiel hierbei auf die Nutzung und Interaktion mit Software aus dem Social Media-Bereich. Konkret haben wir uns für Twitter[6] als Software-System entschieden, da Twitter eine weltweit sehr stark genutzte Social-Media Plattform darstellt [424, 425]. Die technische Umsetzung und Evaluation erfolgte unterstützend durch die Masterarbeit von Felix Volodarskis [6].

7.3.1 Entwicklung und Implementation des Prototyps

Zunächst wurden iterativ Papierprototypen angefertigt, um ein erstes Design-Konzept für die Privacy Explanations auszuloten und zu überlegen, wie diese anschließend in die Twitter-Umgebung eingebettet werden können. Der finale Prototyp bestand aus einer der Twitter-Software identischen Benutzeroberfläche[7], so dass die Teilnehmer bei der späteren Benutzerstudie ein ihnen bekanntes „*Look 'n Feel*“ erhielten, ohne sich an eine neue und unbekannte Benutzeroberfläche gewöhnen zu müssen.

[6] `https://twitter.com`

[7] `https://github.com/CleverProgrammers/twitter-clone`

Unser Prototyp sollte – gemäß dem **3C-Prinzip** (Abschnitt 7.2.2) – mit kontextuellen Privacy Explanations arbeiten. Daher haben wir im Vorfeld verschiedene solcher Kontextszenarien definiert. Diese Kontextszenarien entsprechen realen Interaktionen von Endbenutzern mit der Twitter-Software. Als Szenarien wurden das (i) Posten eines Bildes oder GIFs[8], (ii) die Angabe des eigenen Standorts, (iii) die Erstellung und Interaktion mit Posts sowie (iv) das Suchen von Inhalten gewählt. Für jede dieser Szenarien wurde eine eigene Privacy Explanation gemäß unseren Forschungsergebnissen aus Abschnitt 7.2 entworfen, die wiederum spezifische Informationen zur jeweiligen Datennutzung des zugrunde liegenden Szenarios bereitstellt. Weitere Details zu den Szenarien sind in Abschnitt E.4.1 aufgeführt. Eine grafische Darstellung einer zum Einsatz kommenden Privacy Explanation im Look 'n Feel des Twitter-Klons ist in Abbildung 7.13 abgebildet.

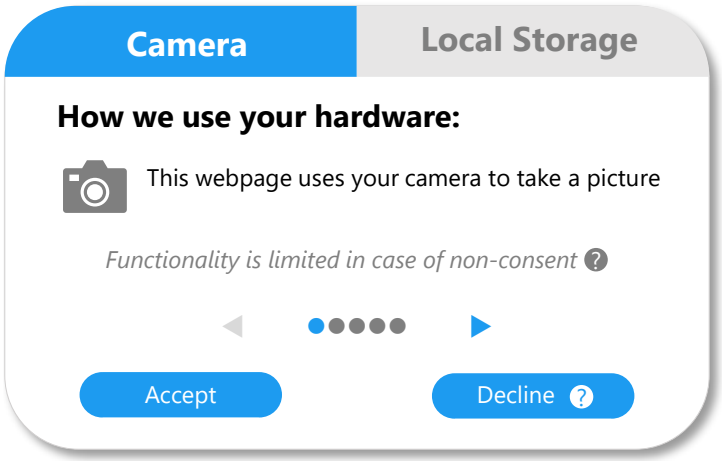


Abbildung 7.13: Prototypische Darstellung einer Privacy Explanation in Anlehnung an [6]

Diese Privacy Explanation wird Szenario (i) entsprechend angezeigt, wenn der Benutzer ein Bild hochladen möchte. Diese kann entweder über seine Kamera oder das Hochladen eines vorhandenen Bildes geschehen. Dargestellt ist hier die Baseline Explanation zur Kameranutzung. Der Benutzer kann durch die einzelnen Erklärungsebenen mit den Pfeilen (links, rechts) navigieren sowie Informationen abrufen, welche Einschränkungen er zu erwarten hat, sollte er der Nutzung der Kamera nicht zustimmen.

[8] GIF steht für Graphics Interchange Format. Ein GIF kann kurze animierte Bilder enthalten und sie werden im Internet häufig als Reaktion oder Kommentar auf einen Post verwendet.

7.3.2 Evaluation

Um das Ziel der Nutzerstudie zu formulieren, haben wir das Goal Definition Template [53, 54] angewandt. Die Formulierung lautete wir folgt:

> ***Goal Definition:*** **Analysiere** das Konzept der kontextuellen Privacy Explanations **in Bezug auf** Nutzbarkeit, Akzeptanz und Verständlichkeit **aus der Sicht von** Endnutzern **im Kontext** einer Online-Studie unter Verwendung der Think-Aloud-Methode.

Das Hauptaugenmerk der Studie besteht somit darin, wie Endbenutzer auf die kontextuellen Privacy Explanations in einem ihnen bekannten Software-System reagieren und wie sich diese auf die Usability auswirken.

7.3.2.1 Design und Durchführung der Studie

Um das oben genannte Ziel zu erreichen und die Aspekte zu analysieren, wurde ein *Synchronous Remote Usability Test* [314] mit einem Online-Fragebogen kombiniert. Um die Qualität der Umfrage zu gewährleisten, wurde sich an etablierte Richtlinien für die Umfrage- und Usability-Testgestaltung gehalten [316, 317, 318]. Wir haben uns für ein Studien-Design mit zwei Gruppen entschieden, da wir Gewöhnungs- und Lerneffekt vermeiden wollten, die die Studienergebnisse beeinflussen könnten. Daher wurden die Teilnehmer nach dem Zufallsprinzip in zwei Gruppen aufgeteilt. Die *Experimentalgruppe*, die mit den kontextuellen Privacy Explanations im Prototyp interagierte und die *Kontrollgruppe*, die die Aufgaben ohne Privacy Explanations zu bewältigen hatte. Beide Gruppen mussten dieselben Aufgaben bewältigen. Ein Überblick über die gestellten Aufgaben, ist in Abschnitt E.4.2 zu finden. Bei der Wahl der Aufgaben wurde darauf geachtet, dass diese tatsächlich realistischen Aufgaben im jeweiligen Kontextszenario entsprachen. Eine Überprüfung des Verständnisses wurde bei den Privacy Explanations zur Kameranutzung und der Auswahl eines GIFs über den Drittanbieter Giphy[9] durchgeführt. Beide Erklärungen sind von komplexerer Natur und die Ergebnisse aus Absatz 7.2.5.4 zeigten, dass die Mehrheit der Benutzer Erklärungen zur Datennutzung von Drittanbietern als am komplexesten in Bezug auf das Verständnis empfanden.

Da es sich um eine Online-Studie handelte, kam Software wie TeamViewer[10] bzw. AnyDesk[11] zum Einsatz, um den Teilnehmern Zugang zu unserem Studien-Computer zu gewähren. Auf diese Weise brauchten die Teilnehmer keine Software oder Plugins auf ihren eigenen Computern installieren. Darüber hinaus wurden Voice Chat-Anwendungen für die Kommunikation

[9] Giphy ist ein Anbieter im Internet, über der Benutzer animierte GIF-Dateien suchen und teilen können.
[10] `https://www.teamviewer.com`
[11] `https://anydesk.com`

genutzt. OBS[12] wurde eingesetzt, um den Bildschirm des Computers des Experimentators aufzuzeichnen und das Experiment anschließend zu analysieren. Während der Studie wurde zudem die *Think-Aloud*-Method [319] angewandt, deren Daten ebenfalls in die anschließende Analyse einflossen.

7.3.2.2 Datenerhebung und Analyse

Die Studie wurde zwischen Januar und Februar 2023 mit insgesamt 62 Teilnehmern durchgeführt. Die Experimentalgruppe bestand ebenso wie die Kontrollgruppe aus 31 Teilnehmern. Die Teilnehmer wurden den jeweiligen Gruppen randomisiert zugewiesen. Die Teilnehmerakquise fand über akademische Kontakte und Anfragen an Studierende der Leibniz Universität Hannover statt. Die Daten wurden sowohl qualitativ als auch quantitativ ausgewertet. Die getesteten Nullhypothesen sind in Tabelle 7.5 aufgelistet.

Tabelle 7.5: Nullhypothesen für die quantitative Datenanalyse

ID	Hypothesen
$H1_0$	Es gibt bezüglich der Durchführungszeiten keinen Unterschied zwischen Experimental- und Kontrollgruppe.
$H2_0$	Es gibt bezüglich des wahrgenommen Verständnisses keinen Unterschied zwischen der Privacy Explanation für Giphy und der für die Kameranutzung.
$H3_0$	Es gibt bezüglich des Verständnisses keinen Zusammenhang zwischen der Lösung der Aufgaben in Bezug auf die Privacy Explanation für Giphy und der für die Kameranutzung.

[12] https://obsproject.com

7.3.2.3 Ergebnisse der Studie

Das Alter der Teilnehmer reichte von 17[13] bis 58 Jahren (M=27,7, SD=8,4). 38,7% waren Studenten und 34.2% der berufstätigen Teilnehmer waren in der IT-Branche tätig. 46.8% der Teilnehmer waren mit Twitter vertraut, davon wurden 14 Teilnehmer der Experimentalgruppe und 15 Teilnehmer der Kontrollgruppe zugewiesen.

Tabelle 7.6: Ergebnisse der Singnifikanztests

ID	Ergebnisse	Signifikant?
$H1_0$	$t \approx 6,1674,\ p < 0,00001$	Ja
$H2_0$	$z \approx -4,5409,\ p < 0,0001$	Ja
$H3_0$	$\chi^2(2,62) = 8,3518,\ p < 0,05$	Ja

Akzeptanz und Nutzbarkeit Es wurde die Durchführungszeit zur Bearbeitung der Aufgaben gemessen und anschließend zwischen den Gruppen verglichen. Die quantitative Analyse (T-Test) lieferte für die Experimentalgruppe signifikant längere Durchführungszeiten ($t \approx 6,1674,\ p < 0,00001$) im Gegensatz zur Kontrollgruppe, so dass $H1_0$ verworfen werden kann (siehe Tabelle 7.6). Der Grund hierfür liegt zum einen in der Natur von Erklärungen im Allgemeinen, da diese von einem Endbenutzer kognitiv verarbeitet werden müssen. Zum anderen werden die Erklärungen im jeweiligen Kontextszenario des Prototyps automatisch angezeigt, so dass der Benutzer mit ihnen in Interaktion treten muss. Die Implementation erfolgte durch Popups, so dass, unabhängig davon ob der Nutzer sich mit der Privacy Explanation mental auseinander setzt oder nicht, diese zumindest „wegklicken“ musste. Auf den ersten Blick ist es wenig überraschend, dass die Experimentalgruppe länger brauchte als die Kontrollgruppe und ist intuitiv auch leicht nachzuvollziehen. Stellt ein System ein zusätzliches Interaktionselement bereit, verlängert das die Interaktionszeit mit diesem System. Interessant ist jetzt aber die Frage, auch wenn der Benutzer mehr Aufwand/Zeit investieren muss, sieht er dennoch einen Mehrwert in den Privacy Explanations? Hier äußerten 17 Probanden der Experimentalgruppe, dass die Privacy Explanations ihren Interaktionsfluss mit dem System beeinträchtigt haben, gleichzeitig gaben aber 82.4% der 17 Teilnehmer an, durchaus einen Mehrwert in den Privacy Explanations für sich empfunden zu haben. Zudem gaben die Teilnehmer an, die Privacy Explanations als *beruhigend* in Bezug auf ihre Privatsphäre empfunden zu haben.

Als weiteres objektives Maß der Bedienbarkeit wurden den Teilnehmern der Experimentalgruppe Fragen des SUS [320] gestellt (siehe Abschnitt E.4.3). Die errechnete SUS-Punktzahl lag im Durchschnitt bei 79,57. Das ist um 11,37 Punkte höher als eine durchschnittliche

[13] Einer der Teilnehmer war minderjährig. Für diese Person lag eine schriftliche Einverständniserklärung eines Erziehungsberechtigten vor.

Webanwendung [325]. Das berechnete 95% Konfidenzintervall $[74,31,\ 84,83]$ zeigt, das der Prototyp mit kontextuellen Privacy Explanations gute bis sehr gute Benutzerfreundlichkeit aufweist [325].

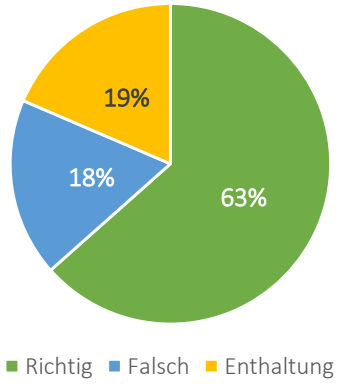


Abbildung 7.14: Überprüftes Verständnis

Verständlichkeit und Relevanz Um einen Eindruck des Benutzerverständnisses der gegebenen Privacy Explanations zu erhalten, wurden die Teilnehmer sowohl nach ihrer Selbsteinschätzung in Bezug auf die Verständlichkeit der Privacy Explanations gefragt als auch gezielt Fragen zum Inhalt der gegebenen Erklärungen gestellt. Die aggregierten Ergebnisse (beide Erklärungen) für das überprüfte Verständnis sind in Abbildung 7.14 dargestellt. Insgesamt konnten die Fragen von 63% der Teilnehmern korrekt beantwortet werden. Wobei hier ein Unterschied zwischen Giphy (56,5% richtig) und der Privacy Explanation zur Kameranutzung (70,3% richtig) zu verzeichnen ist. Der Zusammenhang zwischen den Variablen war hier signifikant (siehe $H3_0$ in Tabelle 7.6). Die Probanden haben die Privacy Explanation zur Kameranutzung mit größerer Wahrscheinlichkeit richtig verstanden.

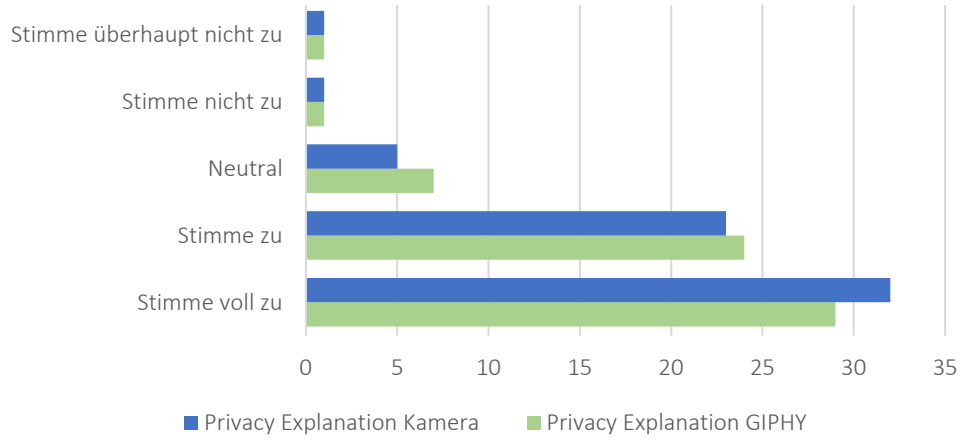


Abbildung 7.15: Überblick der Einflüsse von Privacy Explanations (beide Gruppen)

Darüber hinaus wurden die Teilnehmer vorher nach ihrer Selbsteinschätzung bezüglich des Verständnisses gefragt. Abbildung 7.15 stellt das Ergebnis für die zwei gewählten Privacy Explanations grafisch dar. Festzustellen ist, das die Selbsteinschätzung der Teilnehmer mit den Ergebnissen des Verständnistests in Bezug auf die unterschiedlichen Erklärungen übereinstimmt. Aggregiert man die Zustimmung sowie die Nichtzustimmung der Teilnehmer, glauben 85,5% die Erklärung zu Giphy und 88,7% die Erklärung zur Kameranutzung verstanden zu haben. Die Einschätzung der Teilnehmer steht also im Kontrast zum überprüften inhaltlichen

Verständnis. Zudem weisen beide Verständnistests (subjektive Wahrnehmung und inhaltliches Verständnis) einen signifikanten Unterschied zwischen den Erklärungen auf.
Die Teilnehmer der Studie sollten die für sich empfundene Relevanz der gegebenen Privacy Explanations auf einer Likert-Skala (*Stimme voll zu* bis *Stimme überhaupt nicht zu*) angeben. Das Ergebnis – grafisch abgebildet in Abbildung E.8 – zeigt ein interessantes Bild (siehe Tabelle 7.7). Obwohl die Erklärung zu Giphy für die Teilnehmer nicht ganz verständlich war, stuften sie diese Erklärung als nicht so relevant ein. Als Grund dafür haben die Teilnehmer angegeben, das es sich hier nicht um die eigenen Bilder handele wie zum Beispiel bei der Kamera.

Tabelle 7.7: Wahrgenomme Relevanz der einzelnen Privacy Explanations

Privacy Explanation	$\bar{x}$	**Relevant?**
Kamera	4,13	Ja
Lokaler Speicher (Bild)	4,10	Ja
Giphy	2,73	Nein
Lokaler Speicher (GIF)	3,23	Nein
GPS	4,40	Ja
Manual Explanation (Location)	3,97	Ja
Tweet	3,90	Ja
Suche	3,26	Ja
Interaktion (Tweet)	3,26	Nein
Lesezeichen	3,08	Nein

7.4 Einschränkungen und Bedrohung der Validität

Literaturrecherche. Eine LR erfordert ein gemeinsames Verständnis der von allen Forschern verwendeten Methoden und Konzepte (Such- und Analysemethoden). Die Ergebnisse könnten verfälscht werden, wenn Methoden und Konzepte missverstanden werden. Wir begegneten diesem Threat, indem wir ein Überprüfungsprotokoll erstellten und es zu Beginn der Überprüfung besprachen, um ein ausreichendes gemeinsames Verständnis zu erreichen. Wir formulierten Einschluss- und Ausschlusskriterien, um eine Verzerrung durch subjektive Entscheidungen in unserem Selektionsprozess der Publikationen zu verringern. Die Analyse der Literaturdaten und Validierung wurde eingangs von Jakob Droste und anschließend von mir unabhängig voneinander durchgeführt. Bei Meinungsverschiedenheiten haben wir entsprechende Ergebnisse so lange diskutiert, bis ein Konsens zwischen uns erzielt wurde. **Benutzerstudien**. Obwohl 61 bzw. 62 Teilnehmer eine aussagekräftige Stichprobengröße darstellen, könnten einige der Schlussfolgerungen durch diese Größe beeinflusst werden und sollten daher nicht übergeneralisiert werden. Die Strategien zur Auswahl der Teilnehmer weist einige Einschränkungen auf und spiegelt möglicherweise nicht die gesamte Population wider, was die Verallgemeinerbarkeit unserer Ergebnisse gefährden könnte. Die Mehrheit unserer Teilnehmer verfügte über fundierte Kenntnisse in der Informationstechnologie (IT), d.h., Personen, die Schwierigkeiten mit der Bedienung von Softwaresystemen haben, werden möglicherweise nicht ausreichend berücksichtigt. Wir haben jedoch keine Anzeichen dafür gefunden, dass sich dies auf unsere Ergebnisse auswirkt. Stattdessen haben wir wertvolle Erkenntnisse darüber gewonnen, was verschiedene Menschen denken und welche Einstellung sie zum Datenschutz haben. Die Wahrnehmung und Verständlichkeit von Erklärungen ist schwer zu messen, und es mangelt an geeigneten Metriken [3]. Wir sind mit dieser Gefahr umgegangen, indem wir statistische Tests verwendet haben, wo es angebracht war, und ansonsten eher qualitative Analysen durchgeführt haben.

7.5 Fazit

Wir haben gezeigt, welche Informationen Endbenutzer in Bezug auf die Verwendung ihrer Privatsphäreaspekte als relevant zu erachten scheinen. Dazu zählt nicht nur *das* ein Privatsphäreaspekt genutzt wird und *wofür*, sondern auch wofür dieser *nicht* genutzt wird. Erreicht wird die Kommunikation der jeweiligen Datenpraktiken über verschiedene Arten von Erklärungstypen, wobei die Contrastive Explanation die **Nichtnutzung** umfasst. Eine Herausforderung beim Design im Umgang mit den Contrastive Explanations zeigte sich im

Prototyp der Konzeptstudie bereits bestätigt. Die Erklärungen zur Nichtnutzung von Privatsphäreaspekten wird teilweise von Benutzern als zu kurz wahrgenommen. Das liegt vor allem daran, das hier die zu kommunizierenden Informationen beliebig „tief“ ausgerollt werden können. Nutzt ein System beispielsweise den Standort eines Nutzers zum Herausfinden der nächsten Bushaltestelle, könnte eine potentielle Contrastive Explanation Nichtnutzung über „*Weitergabe der Daten an Dritte*“, „*persistente Speicherung der GPS-Daten*“, „*Verknüpfung des Standorts mit einem Benutzerkonto, um Bewegungsprofile zu erstellen*“ enthalten, aber auch Dinge wie „*Nichtermittlung der nächsten Bahnhaltestelle*“, „*Bestimmung des Wetters am Standort*“ etc. Also prinzipiell kann diese Liste mit mehr oder weniger relevanten Informationen beliebig weitergeführt werden. Daher ist es essentiell, bereits in der Base-line Explanation die Verwendung des Privatsphäreaspekts möglichst präzise zu formulieren und einzugrenzen.

Dass die Example-Based Explanation im Kontext von Privacy Explanations als sehr verständlich empfunden wird, rechtfertigt ihre gute Eignung als Erklärungstyp. Im Kontrast dazu steht die geringer wahrgenommene Relevanz dieses Erklärungstyps in der Konzeptstudie. In der Studie des technischen Prototyps hingegen wurde sie als sehr relevante Erklärung empfunden. Dieses Phänomen lässt sich möglicherweise damit erklären, das die Example-Based Explanation als eine Art Basisanforderung, wie im Kano-Modell[14] beschrieben, angesehen werden kann. Basisanforderungen haben eine hohe Erfüllungspflicht und beschreiben, „das Minimum, was ein System leisten muss“, damit überhaupt ein Grundmaß an Kundenzufriedenheit erreicht werden kann. Häufig werden diese Anforderungen aber nicht explizit kommuniziert, da der Kunde diese als selbstverständlich ansieht. Ähnlich könnte das mit dem Erklärungstyp der Example-Based Explanation sein. Dieser transportiert auf sehr verständliche Weise notwendig zu kommunizierende Informationen. Diese werden aber vom Benutzer als selbstverständlich wahrgenommen und nehmen daher in der persönlichen Relevanzgewichtung eine eher untergeordnete Rolle ein.

Die Umsetzung der kontextuellen Privacy Explanations im Prototyp lieferte weitere wertvolle Erkenntnisse. Sie wirken sich zwar negativ auf die Performanz in Bezug auf die Interaktionsgeschwindigkeit mit einem System aus, werden mehrheitlich von den Benutzern aber dennoch als nützliches Feature angesehen. In der Literatur ist dieses Phänomen der Erklärbarkeit als Double-Edged-Sword-Effekt bekannt [165]. Erklärbarkeit kann, wie andere Qualitätsaspekte auch, antagonistische Beziehungen mit anderen Qualitätsaspekten haben. Ein prominentes Beispiel hierfür ist die Beziehung von Usability und Security. Die Passwortabfrage in einem Software-System wird i.d.R. als „nervig“ und „störend“ empfunden [426, 427], ist aber ein notwendiges und gewolltes Sicherheitsmerkmal.

[14] Grundlegende Informationen zum Kano-Modell sind im Anhang in Abschnitt A.1 zu finden.

Sowohl im Prototyp der Konzeptstudie als auch im technischen Prototyp waren die Third Party Explanations herausfordernd, was deren Verständnis und Teilnehmer-Feedback anging. Bei Durchsicht der Teilnehmerdaten wird deutlich, dass den Benutzern hier häufig zu wenig Informationen bereitgestellt werden. In der Konzeptstudie ging es in erster Linie um Wahrnehmung dieser Information hinsichtlich Relevanz und Informiertheit. Die Umsetzung im technischen Prototyp zeigte dann aber, das hier noch weiterer Forschungsbedarf in Bezug auf die Umsetzung der Third Party Explanation nötig ist, denn die bloße Nennung des Unternehmens, das ggf. auch Zugriff auf Daten erhalten hat mit samt einem Link zu deren DSE, ist möglicherweise nicht ausreichend.

Beantwortung RQ3: Privacy Explanations können das Vertrauen in ein Software-System fördern, da sie den Endnutzern ein Gefühl der Sicherheit vermitteln können, für mehr Transparenz in Bezug auf die angewandten Datenpraktiken eines Systems sorgen, die Endnutzer in die Lage versetzen, bewusstere Entscheidungen (informed consent) hinsichtlich ihrer Privatsphäre zu treffen und zudem das Bewusstsein für den Datenschutz (Privacy Awareness) fördern können. Diese Faktoren können sich ebenfalls auf die erlebte Vertrauenswürdigkeit eines Softwaresystems auswirken. Endnutzer möchten gemäß dem *2W1H-Prinzip* über Datenschutzfragen informiert werden. Sie erwarten auch zu erfahren, ob ihnen Nachteile entstehen, wenn sie einer möglichen Datenverwendung nicht zustimmen. Die Bereitstellung dieser relevanten Informationen im in Bezug stehendem Nutzungskontext (*3C-Prinzip*) versetzt die Nutzer in die Lage, ihr Recht auf Selbstbestimmung im Rahmen ihrer Privatsphäre zu wahren und aktiv eingreifen (zustimmen/ablehnen) zu können. Was das Design der Privacy Explanations anbelangt, ist der Einsatz unterschiedlicher Erklärungstypen kombiniert mit einem Layered-Ansatz eine geeignete Strategie, Endbenutzern die Verwendung von Privatsphäreaspekten auf unterschiedlichen Granularitätsleveln, entsprechend ihrer *Privacy Attitude*, zu präsentieren.

8

Conclusio

Bei der Nutzung digitaler Systeme generieren Nutzer eine enorme Menge an Daten und hinterlassen dadurch Spuren ihres „Online-Selbst“, so dass vieles, was wir online oder offline tun oder sagen, nicht mehr *flüchtig* ist [428]. All diese Daten werden für gewöhnlich gespeichert, zusammengeführt und ausgewertet [26, 30, 31]. Beispielsweise enthält ein digitales Foto eine Reihe an Metainformationen wie Zeitstempel, Standort (GPS-Koordinaten), Kamerainformationen und Einstellungen [31], die viel über unser persönliches und soziales Umfeld preisgeben können [429]. Die Verarbeitung oder Weitergabe der persönlichen Daten geschieht häufig ohne Wissen oder explizite Zustimmung der Nutzer. Die i.d.R. einzige Quelle für Endbenutzer, um Informationen zu Datenpraktiken eines Service-Anbieters zu erhalten sind Datenschutzerklärungen (DSEs). Diese haben sich allerdings für Endbenutzer in ihrer derzeitigen Form als nicht tauglich erwiesen [43, 46, 47] (siehe auch Kapitel 5). Es mangelt an einer nutzerzentrierten Lösung, die Endbenutzer auf verständliche Weise über Datenpraktiken informiert und aufklärt, so dass diese möglichst in der Lage sind, bewusste Entscheidungen in Bezug auf ihre Privatsphäre zu treffen. Diese Kluft der Informationsasymmetrie – zwischen Endbenutzer und digitalem System – versucht meine Dissertation zu überbrücken und Lösungen für dieses Problem vorzuschlagen.

8.1 Grenzen der Arbeit und zukünftige Forschungen

Die Privacy Explanations wurden in einer Konzeptstudie evaluiert und prototypisch in ein bestehendes Software-System integriert. Beide Benutzerstudien umfassten jeweils eine Teilnehmerzahl von N > 60. Um weitere Erkenntnisse zu gewinnen sowie aussagekräftigere Ergebnisse hinsichtlich Usability und Privacy Awareness zu generieren, sollte der Teilnehmerkreis deutlich erweitert werden und die Studie über einen längeren Zeitraum stattfinden.

Sowohl in der Konzeptstudie als auch im technischen Prototyp merkten die Teilnehmer an, dass sie gerne weitere Information zu den Datenpraktiken der Drittanbieter hätten. Dies sollte im Zuge zukünftiger Forschungen ebenfalls untersucht werden, wie diese Informationen kommuniziert werden können und woher diese stammen könnten.

Ein wichtiger Faktor, der die *trustworthiness* der Privacy Explanations betrifft, ist die juristische Verbindlichkeit der Informationen. Hierfür empfehlen sich interdisziplinäre Forschungen zwischen Software Engineers und Juristen, um auszuloten, wie die Beschaffenheit der Erklärungen auszusehen hat, um *(a)* Gesetzen und Regulierungen zu entsprechen, *(b)* garantierte Rechtsverbindlichkeit für den Benutzer zu schaffen und *(c)* die Einfachheit und Verständlichkeit der kommunizierten Informationen zu gewährleisten.

Die Perspektive der Industrie sollte zukünftig unbedingt berücksichtigt werden. Hier geht es darum herauszufinden, wie sich das Konzept der Privacy Explanations in kommerziellen Produkten implementieren lässt. Dabei ist es wichtig, mögliche Bedenken von Seiten der Industrie wie auch Chancen zu diskutieren und gegeneinander abzuwägen.

8.2 Zusammenfassung der Ergebnisse

Für die im Rahmen dieser Dissertation entstandene Forschung habe ich drei Theorien entwickelt, welche sich auf sieben Artefakte stützen. Die Artefakte dienten wiederum der Beantwortung der drei Forschungsfragen (RQs) und adressieren somit die identifizierte Forschungslücke (relevantes Problem). Die geschaffenen Artefakte sollen zum einen Endbenutzer bezüglich ihrer Privatsphäre informieren und aufklären sowie auch als Ausgangspunkt für Software Engineers dienen, Privacy Explanations im Software-System zu integrieren. Die relevantesten Erkenntnisse meiner Arbeit sind, **(i)** dass Zusammenspiel von Erklärbarkeit und Privacy zur Erreichung weiterer Qualitätsziele führen kann und diese positiv beeinflusst (Abschnitt 8.2.1), **(ii)** das Privacy Explanations sich als nutzerzentrierte Lösung zur Kommunikation von Datenpraktiken eignen (Abschnitt 8.2.2) und **(iii)** das die geschaffenen Artefakte beim Design von Software-Systemen als probates Mittel hin zu mehr privatsphärefreundlichen

Systemen dienen können (Abschnitt 8.2.3). Nachfolgend werden die genannten Erkenntnisse etwas tiefer gehend beleuchtet.

8.2.1 Privatsphäre und das Zusammenspiel mit Erklärbarkeit

Security und Privacy sind fundamental miteinander verwoben. „Wenn wir keine Privatsphäre haben, fühlen wir uns ungeschützt und verletzlich; wir fühlen uns weniger sicher. Ähnlich verhält es sich, wenn unsere persönlichen Räume und Daten nicht sicher sind: Wir haben weniger Privatsphäre“ [31, S.156]. Ein Blick auf die „Security versus Privacy“-Debatte [430, 431] offenbart hierzu folgenden oft angeführten Trade-Off [432, 433]: „für mehr Sicherheit müssen wir unsere Privatsphäre opfern und uns der Überwachung unterwerfen. Und wenn wir ein gewisses Maß an Privatsphäre wollen, müssen wir erkennen, dass wir dafür etwas Sicherheit opfern müssen“ [31, S.156]. Stimmt das wirklich so? Schaut man hier etwas genauer hin, sieht man schnell, dass dies ein falscher Trade-Off ist und er gar nicht so schwarz weiß ist, wie er vordergründig zu sein scheint und uns die Befürworter beider Lager glaubhaft machen wollen [432, 434, 435]. Die oben angesprochene Verwobenheit von Security und Privacy macht das bereits deutlich und nicht zuletzt hängt es beispielsweise von der Wahl wie Security implementiert wird ab. Türschlösser, Sicherheitspersonal sind Beispiele für Sicherheitsmaßnahmen, ohne die Privatsphäre zu verletzten. Privatsphäre als normatives Konzept schafft den Rahmen für Entscheidungen. Entscheidungen darüber, „wer legitimerweise die Möglichkeit haben sollte, auf Informationen zuzugreifen und sie zu verändern“ [431]. Security implementiert diese Entscheidungen dann in den Systemen. Security dient als eine Art Mediator, der „zwischen den Entscheidungen über Informationen und Privatsphäre“ vermittelt [431].

Der ehemalige Vorstandsvorsitzende von Google, Eric Schmidt, sagte: „Wenn Sie etwas haben, von dem Sie nicht wollen, dass es jemand weiß, sollten Sie es vielleicht gar nicht erst tun“ [436]. Mutmaßliche Rechtfertigungen dieser Art oder denen von Befürwortern der (Massen-) Überwachung: „Wenn man nichts Falsches tut, braucht man auch nichts zu verbergen.“ verzerren immer wieder das Bild bzw. das Konzept der Privatsphäre. Das Problem mit dieser Art von Aussagen ist, dass sie implizit besagen, dass Privatsphäre etwas mit dem Verbergen von etwas Falschem oder Verbotenem zu tun hat. Im Gegenteil, es geht hier nicht um das Verschleiern von irgendwelchen Machenschaften. Wie ich in Abschnitt 6.1 bereits diskutiert habe, ist Privatsphäre wichtig für uns als Individuum. Das Fehlen von Privatsphäre wäre nicht gesund für ein einzelnes Individuum – und auch nicht für eine Gesellschaft.

„Transparenz ist für eine offene und freie Gesellschaft unerlässlich“ [31]. Die Öffnung der Regierung und Verwaltung gegenüber seinen Bürgern sowie die Informationsfreiheit verschaffen diesen Wissen darüber, was ihre Regierung tut. Ähnlich sieht es bei der Offenheit von Unternehmen aus. Je offener ein Unternehmen oder eine Regierung sind, desto besser lässt sich

entscheiden, wie vertrauenswürdig diese sind. Das Gleiche gilt für Software-Systeme, wenn es um unsere persönlichen Daten geht. Der Schutz der Privatsphäre „ist mit einer Reihe von rechtlichen Anforderungen und bewährten Praktiken im Umgang mit personenbezogenen Daten verbunden, wie z. B. der Notwendigkeit, den Verbraucher zum Zeitpunkt der Zustimmung darüber zu informieren, welche Daten erhoben und wie sie verwendet werden“ [271]. Diese Arten der Transparenz können das Vertrauen in Regierungen, Unternehmen und schlussendlich auch Software stärken. Denn „um Vertrauen aufzubauen, muss man Anreize schaffen, damit Vertrauensnehmer [(Organisationen oder Service-Anbieter)] Maßnahmen zur Vertrauenswürdigkeit ergreifen und die Vertrauensgeber [(Endbenutzer)] entsprechende verlässliche Signale empfangen“ [437].

Vertrauen in die Fähigkeit eines Systems bedeutet, dass ein System über die Funktionalität oder die funktionale Fähigkeit verfügt, eine bestimmte Aufgabe zu erfüllen, die der Vertrauensgeber verlangt [438]. In Bezug auf den Datenschutz können undurchsichtige, unverständliche Datenpraktiken, also mangelnde Transparenz, dieses Vertrauen beeinträchtigen. Um diese Beziehung von Vertrauen und Vertrauenswürdigkeit zu verdeutlichen, nutze ich die Metapher einer Münze: einer *Vertrauensmünze*. Auf der einen Seite der Münze steht das Vertrauen des Endnutzers, das sein empfundenes Vertrauen (englisch: *perceived trust*) in ein System darstellt. Auf der anderen Seite der Münze steht die Vertrauenswürdigkeit als Eigenschaft oder Qualitätsaspekt dieses Systems. Im Hinblick auf den Datenschutz können wir die beiden Seiten nicht getrennt betrachten, da sie miteinander verwoben sind. „Die Anbieter (= Vertrauensnehmer) sollen vertrauensbildende Maßnahmen ergreifen und für die Nutzer (= Vertrauensgeber) kenntlich machen, damit sie die Vertrauenswürdigkeit beurteilen und sich entsprechend verhalten können“ [437]. Das bedeutet, wenn Software Engineers wollen, dass die Endnutzer ihrem System vertrauen, sollten sie sicherstellen, dass ihr System in Bezug auf die Privatsphäre vertrauenswürdig ist.

Erklärbarkeit als NFR kann hier als Vermittler dienen. Ein Vermittler, der Transparenz herstellt, z.B. bei der Offenlegung von Datenpraktiken, hier für Verständnis und Aufklärung sorgt sowie Vertrauen und Vertrauenswürdigkeit in und von Software fördern kann. Zwischen Erklärbarkeit und Privacy besteht ein sehr fruchtbares Zusammenspiel, da viele weitere Qualitätsziele in der Software selbst erreicht werden können (siehe Abbildung 4.4), sich aber auch regulatorische Anforderungen, wie sie z.B. der DSGVO entspringen, aber auch Business Goals wie Kundenzufriedenheit oder Nutzerakzeptanz stärken lassen [184, 288]. Ich sehe im Konzept der Erklärbarkeit einen starken Partner für die Privacy im digitalen Zeitalter und hier besonders als Wegbereiter für transparenten, respektvollen Umgang mit den Daten der Nutzer von Seiten zukünftiger Systeme.

8.2.2 Privacy Explanations als nutzerzentrierte Lösung

Es gibt eine Reihe unterschiedlicher Ansätze und Lösungen, DSEs zu „verbessern“ und diese dem Endbenutzer zugänglicher zu machen. Dabei gibt es neben der maschinellen [42, 177, 178, 179, 183] und visuellen Aufbereitung [184, 189, 174] der DSEs auch Forschungen über den Einsatz von Privatsphäresiegeln [195, 196] oder Kennzeichnungen [193, 194], ähnlich den Produktkennzeichnungen von Lebensmitteln oder Elektrogeräten.

Allerdings wird hier versucht ein Informationsmedium, deren Adressaten eigentlich nicht die Endbenutzer selber sind, auf diese umzumünzen, ohne hierbei Aspekte wie Privacy Attitudes, unterschiedliche Nutzerpräferenzen oder nicht so IT-erfahrene Anwender zu berücksichtigen. Darüber hinaus bietet nicht jedes digitale Informationssystem eine DSE oder Vergleichbares und wenn vorhanden, sind diese meist nicht in das System integriert, so dass die Wahrscheinlichkeit sehr gering ist, dass Benutzer diese überhaupt lesen [439]. Ein weiterer sehr gewichtiger Nachteil dieser Herangehensweisen ist, dass hier *Dritte* versuchen, Missstände in der Kommunikation von Datenpraktiken von anderen zu beheben, wo diese doch eigentlich selbst in der Pflicht sind, die Informationen ausreichend und verständlich an den Endbenutzer zu verteilen. Wie es ja auch die verschiedenen Regularien, wie DSGVO, CPRA, FIPPs, etc. vorsehen. Iachelle und Hong [37] sehen eine mögliche Lösung in der „Entwicklung effektiverer und weniger aufwändiger User Interfaces (UIs), die den Menschen helfen, gute Entscheidungen zu treffen“. Also das Problem direkt zu adressieren, ohne Umwege über Dritte. Die Autoren geben aber auch zu bedenken, dass hier sehr behutsam vorgegangen werden muss, da entsprechende UIs und Interaktions-Patterns schnell sehr komplex werden können und die Erfahrung zeigt, „dass die meisten betroffenen Personen nicht in der Lage oder nicht willens sind, alle Offenlegungen personenbezogener Daten zu kontrollieren und den Überblick über alle Parteien zu behalten, die ihre personenbezogenen Daten verarbeiten“ [37, S.97].

Zur Adressierung dieses Problems habe ich das Artefakt der **Privacy Explanations** geschaffen (siehe Abschnitt 6.1.3), um den Nutzer über Datenpraktiken eines digitalen Informationssystems zu informieren. Dieses Artefakt folgt dem *3C-Prinzip* (Abschnitt 7.2.2), stellt dem Benutzer also nur kontextbezogene Informationen zu Privatsphäreaspekten bereit und verbirgt die mögliche Komplexität mit Hilfe des *Layered-Ansatzes* (Abschnitt 7.2.3.2) sowie unterschiedlichen Arten von Erklärungen, um somit möglichst minimal invasiv in Bezug auf die Usability eines Systems zu sein und gleichzeitig den verschiedenen Privacy Attitudes der Benutzer Rechnung zu tragen. Außerdem sind Privacy Explanations eine Komponente des Systems selbst und nicht von diesem losgelöst, so dass der Nutzer automatisch mit ihnen in Interaktion tritt.

Als nutzerzentrierte Lösung stellen Privacy Explanations den Vertrauensgeber (Nutzer), um dessen persönliche Daten es geht, in den Fokus. Das System tritt mit ihm in einen Dialog über

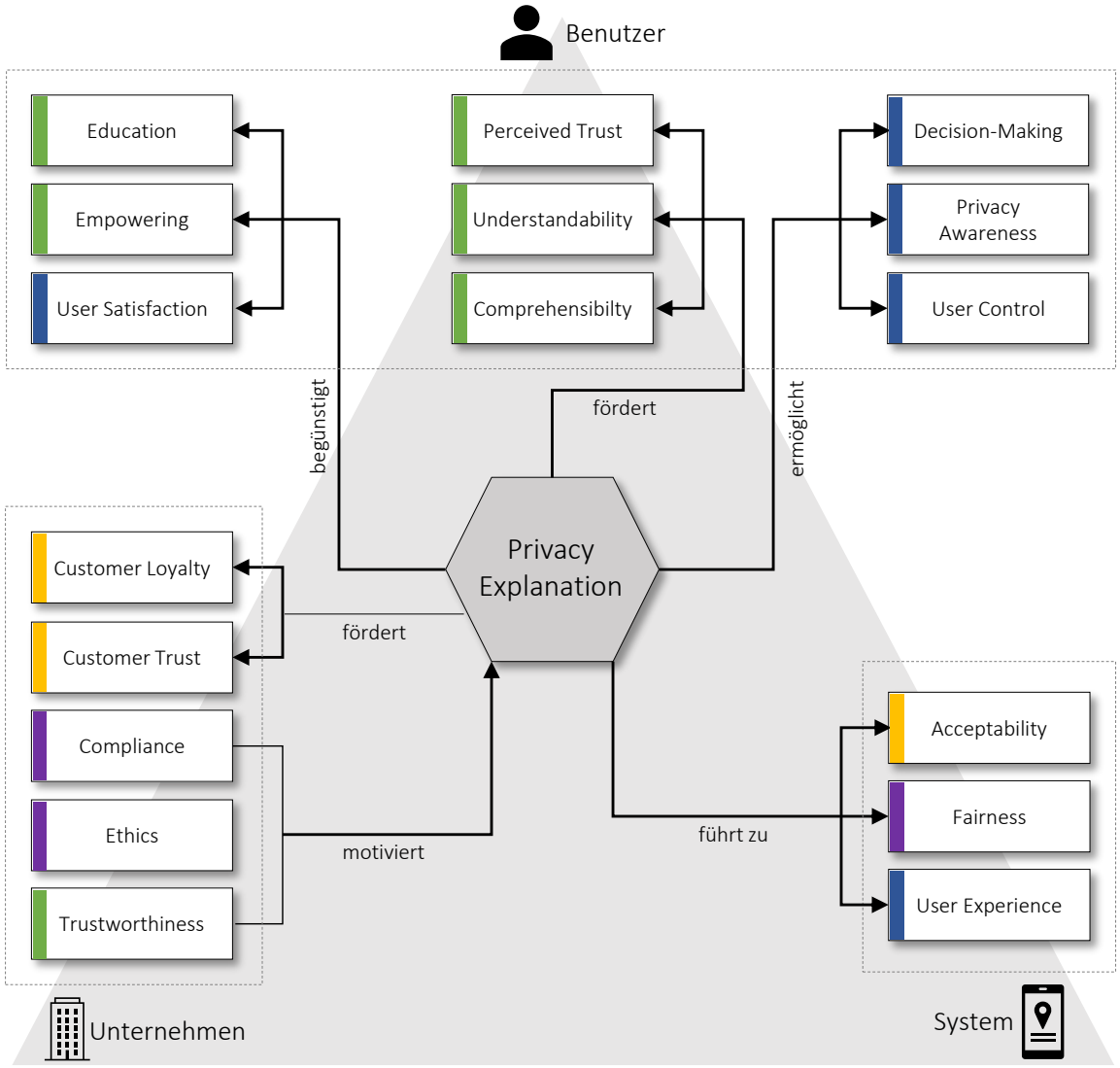


Abbildung 8.1: Modell des Beziehungsdreiecks von Privacy Explanations

die Verarbeitung seiner Daten. Das abgebildete Modell in Abbildung 8.1 soll die Beziehung zwischen Endbenutzer, System und Unternehmen veranschaulichen. Gemäß dem Modell der Beziehung von Erklärbarkeit und Privacy (Abbildung 4.4) existieren noch weitere relevante Qualitätsaspekte, die hier nicht explizit aufgelistet sind. Ich habe mich aus Gründen der Übersichtlichkeit auf die aus meiner Sicht relevanten *Schlüsselqualitäten* beschränkt.

Die Mitteilung und Auskunft über Datenpraktiken stellt das Fundament dar, auf denen weitere wie auch in den FIPPs verankerte Prinzipien wie Wahlmöglichkeit, Zugang zu Informationen, Sicherheit und Umsetzung (der Leitlinien oder Regularien) gründen [270]. Privacy Explanations sollten demzufolge eine Motivation der Unternehmen darstellen, vertrauenswürdige Aufklärung zu leisten, entsprechend der Gesetze und Regularien. Im Gegenzug kann das

auf Seiten des Systems zu mehr Akzeptanz, Fairness und somit mehr User Experience (UX) führen sowie im Gegenzug für das Unternehmen die Kundenbindung und das -vertrauen fördern. Es profitiert also nicht nur der an der Spitze stehende Benutzer, sondern Unternehmen können mit ihren Software-Produkten ebenfalls profitieren. Bestehende Prozesse und Frameworks für den Aufbau datenschutzfreundlicher Systeme, wie beispielsweise Privacy by Design (PbD) [9], konzentrieren sich primär auf die Analyse der Datenpraktiken eines Systems und weniger auf die Gestaltung von Mitteilungen oder Hinweisen zu diesen [184]. Privacy Explanations haben das Potenzial diese Lücke zu schließen. Sie können die Beziehung zwischen Benutzer, System und Unternehmen positiv beeinflussen und können ebenfalls einen Einfluss auf das System-Design haben, was wiederum zu einem ethischeren Umgang mit den Daten der Benutzer führen und zum anderen Datensparsamkeit des Systems fördern kann. Gründe hierfür soll der nächste Abschnitt erläutern.

8.2.3 Privacy Aware Systems Design

Um das Vertrauen der Endnutzer zu gewinnen, müssen die Systeme verlässlich sein, d.h. vertrauenswürdig (englisch: *trustworthy*) in Bezug auf die Privatsphäre. Daher werden Systeme benötigt, die Qualitätsaspekte wie Accountability, Fairness und Ethik erfüllen und bei denen Privacy eine Art „*default setting*“ ist [9]. Zudem müssen diese Systeme die geltenden Gesetze und Vorschriften in Bezug auf den Datenschutz umsetzen, um die Privatsphäre des Einzelnen bei der Informationsverarbeitung zu schützen und zu respektieren [440].

In Bezug auf Privacy Explanations bedeutet dies, dass ein System, das einem Benutzer erklärt, zu welchem Zweck ein bestimmter Privatsphäreaspekt verarbeitet wird, diesen nicht für andere Zwecke verwenden darf. Dies muss das System garantieren, um als vertrauenswürdig zu gelten. Indem es Erklärungen abgibt, kann das System den Endnutzer dabei unterstützen, Vertrauen in das System selbst aufzubauen (siehe Abbildung 8.1). Dies verdeutlicht auch, warum es wichtig ist, nicht nur zwischen *trust* und *trustworthiness* zu unterscheiden, sondern auch beide Qualitätsaspekte gezielt zu adressieren. Die Offenheit oder Transparenz mit der ein System seine Datenpraktiken darlegt, kann zur Fairness beitragen.

Requirements Engineering (RE) und UX Design verfügen über jahrzehntelange Erfahrung im Umgang mit Anforderungen aus den Bereichen Security und Privacy sowie deren Umsetzung und Implementation [441]. Allerdings „stellen die Änderungen der internationalen Datenschutzbestimmungen und die digitale Transformation diese Disziplinen vor neue Herausforderungen“, so Groen et al. [441]. Besonders herausfordernd ist hierbei die Fragestellung, wie zwischen Dateneigentümern, -sammlern und -nutzern Transparenz geschaffen, Entscheidungshilfen für Endbenutzer gegeben, und Bevollmächtigungen erteilt werden können. Diese drei Stakeholder-Gruppen verfolgen unterschiedliche Ziele und resultierend daraus

ergeben sich unterschiedliche, sich möglicherweise im Konflikt befindliche Anforderungen, die in Einklang gebracht werden müssen, um den Stakeholdern gerecht zu werden.

Auch Konzepte wie das Privacy by Design (PbD) oder die ISO 29100 [132] geben zwar mögliche Hilfestellungen zum Design solcher Systeme, aber politischen Entscheidungsträgern, Behörden und der Industrie ist aber derzeit unklar, was PbD denn praktisch bedeutet, damit es funktioniert [332]. Ohm [442] sieht hier ein Problem in der für ihn existierenden Kluft zwischen Technik und Politik. Beide reagieren nur aufeinander, anstatt miteinander zu sprechen. Ohm meint dazu: „Privacy erfordert einen Dialog zwischen zwei Arten von Menschen: denjenigen, die über Richtlinien und Regeln sprechen, und denjenigen, die die Technik erschaffen“ [442]. Das bedeutet, dass politische Entscheidungsträger (z.B. Gesetzgeber und Juristen) in direktem Dialog mit Software Experten zusammenarbeiten und sich gegenseitig dabei unterstützen sollten, Datenschutzanforderungen zu ermitteln, sicherzustellen, dass sie rechtskonform sind, und sie in Systeme, Vorschriften und Normen zu übersetzen [41], die für alle Beteiligten zielführend sind.

Was allerdings in Ohms Aussage fehlt ist der Endbenutzer, der Teil der Beziehung zwischen Dateneigentümern, -sammlern und -nutzern ist. Dieser sollte im Fokus stehen, was in Verbindung mit den Privacy Explanations zum **Trialog der Privacy** führt, der in Abbildung 8.2 zusätzlich grafisch veranschaulicht ist.

> **Trialog der Privacy:** Privacy erfordert einen Trialog zwischen drei Arten von Beteiligten: denjenigen, die Endnutzer sind und deren Privatsphäre betroffen ist, denjenigen, die Privacy Engineering „sprechen“, und denjenigen, die über Regeln und Richtlinien sprechen.

„Privacy lässt sich nicht mit ein paar Architekturentscheidungen in der Design-Phase [eines Systems] allein regeln. Das Engagement für den Datenschutz ist eine ständige Aufgabe […]“ [440]. Das könnte auch eines der Probleme sein, warum Konzepte wie PbD nicht eins zu eins anwendbar sind. Aus diesem Grunde übernehme ich für den Trialog der Privacy, das Konzept der Rolle des Privacy Engineers wie Bowman et al. [440] ihn vorschlagen. Ähnlich zum Usability Engineer, kümmert sich der Privacy Engineer über den gesamten Lebenszyklus eines Systems darum, dass es *privacy-aware* ist und bleibt. Während der Usability Engineer in die UX-Prozesse eingebunden ist und diese mitgestaltet sowie dafür sorgt, dass die Unternehmenswerte hinsichtlich der UX umgesetzt werden, stellt der Privacy Engineer sicher, dass das Produkt so entwickelt, hergestellt und verwendet wird, das es mit den Unternehmenswerten in Bezug auf den Datenschutz konform ist [440]. Dabei ist er im Austausch mit Policy Makers und setzt sich ebenso für Bedarfe der Endbenutzer in Bezug auf deren Privatsphäre ein. So ist es zwischen Technik und Politik kein Vorangehen und Folgen bzw. Reagieren mehr, sondern ein gemeinsames Gestalten in gegenseitigem Dialog.

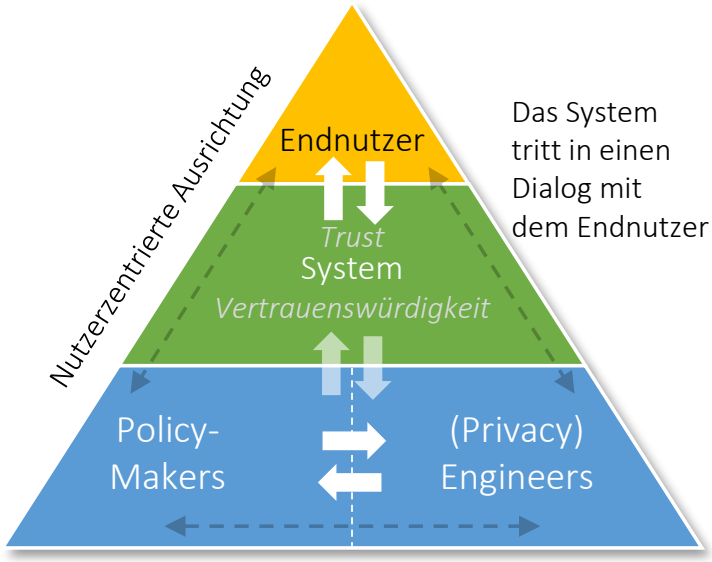


Abbildung 8.2: Trialog der Privacy

Am Beispiel der Privacy Explanations könnte der Privacy Engineer im Dialog mit Juristen und politischen Entscheidungsträgern Konzepte für die Rechtsverbindlichkeit dieser und die Wahrung der Privatsphäre der Endbenutzer ausarbeiten. Mit diesem Wissen dann anschließend den Design-Prozess unterstützen, so dass das System später in einen Dialog mit den Endbenutzer treten kann.

Natürlich kann man auf der Gegenseite argumentieren, dass es zusätzlich zu den vielen verschiedenen bereits existierenden Rollen im Software-Entwicklungsprozess wenig Sinn macht (Kosten, Manpower, Unternehmenskultur, etc.) eine weitere Rolle einzuführen. Dennoch ist es wichtig für Unternehmen und zukünftige Technologien, Software bereits im Vorfeld rechtssicher und datenschutzkonform zu gestalten. Gründe hierfür wurden bereits ausreichend diskutiert. Um die Software aber im Vorfeld derart zu gestalten ist es notwendig zu wissen, was die Regularien besagen und wie diese in Software umzusetzen sind. Auf der anderen Seite könnten Gesetze und Regelungen von vornherein mit dem technisch notwendigen Wissen angereichert sein, anstatt der Technologie immer einen Schritt hinterher zu sein, um wiederholt nachbessern zu müssen. Der Dialog von Policy Makers und Engineers, mit Ausrichtung auf den Endbenutzer, kann hier für Gleichgewicht sorgen.

Zwischen den Domänen des Usability- und Software Engineerings bestand anfangs ebenfalls eine Kluft. Ein Grund war damals gewesen, das die User-Centered Design (UCD)-Techniken, die von der Human-Computer Interaction (HCI) Community entworfen wurden, nicht auf bekannte Methoden und Techniken des SE zurückgriffen und der Einsatz der neuen Methoden unverständlich war [443]. Heute sind SE und UCD eng miteinander verflochten, aber nicht jedes Entwicklungsteam leistet sich einen eigenen Usability Engineer. Das gleiche

würde für den Privacy Engineer gelten. Kleinere Teams könnten auf Techniken oder Methoden zurückgreifen, wie beispielsweise existierende Privacy Patterns[1]. Hierbei handelt es sich um Lösungen für verschiedene Privacy-Probleme, die sich am PbD orientieren. Es könnte auch ein Privacy Engineer für mehrere Teams innerhalb eines Unternehmens zuständig sein sowie auch Schulungen angeboten werden und auch ein Modell als externer Consultant (Unternehmensberater), wie es z.B. bei Requirements Engineers der Fall ist, wäre denkbar.

Die richtige Antwort wie viel Privatsphäre jeder haben sollte bzw. wie der optimale Schutz von Privatsphäre in einem Informationssystem ausschaut, gibt es vermutlich nicht. „In Anbetracht ihrer Macht als Akteure des Wandels [...], haben [Software] Engineers eine Verantwortung gegenüber dem Rest der Gesellschaft“ [440, S.17]. Daher, so Bowman et al. [440] weiter, sei es inakzeptabel für Software Engineers, „eine agnostische Sichtweise einzunehmen“. Ihre Rolle sollte darin bestehen, sich aktiv in den Trialog zu begeben und gemeinsam mit anderen Disziplinen zu debattieren, wie die Zukunft der Informationssysteme gestaltet werden soll unter Berücksichtigung gesellschaftlich fairer und ethischer Normen, denn „in einer freiheitlich-demokratischen Gesellschaft muss die soziale Verantwortlichkeit in Bezug auf die Privatsphäre Teil der technologischen Entwicklung sein“ [440, S.17].

[1] https://privacypatterns.org

A Grundlagen

A.1 Grundlagen Kano-Modell

Die nachfolgenden Grundlageninformationen sind meiner Masterarbeit entnommen [444] und dort bereits veröffentlicht. „Das Kano-Modell zeigt den Zusammenhang zwischen der Kundenzufriedenheit und der Erfüllung von Kundenanforderungen“ [445]. Die Kundenanforderungen werden hierbei in drei Gruppen unterteilt. Die **Basisanforderungen**, **Leistungsanforderungen** und die **Begeisterungsanforderungen** wie in Abbildung A.1 abgebildet.

Die Basisanforderungen haben eine hohe Erfüllungspflicht in einem Produkt. Ohne diese kann es keine Kundenzufriedenheit geben. Jedoch wird die alleinige Berücksichtigung der Basisanforderungen nicht zur vollen Zufriedenheit des Kunden beitragen. Wie im Modell zu sehen, nähern sie sich asymptotisch der „Erwartung erfüllt“-Achse. D.h. immer mehr erfüllte Basisanforderungen steigern die Kundenzufriedenheit nicht immer weiter. Zudem ist die Erhebung dieser Anforderungen nicht ganz einfach, da der Kunde diese möglicherweise als selbstverständlich ansieht.

Zur Erreichung der Kundenzufriedenheit sind die Leistungsanforderungen notwendig. Im Kano-Modell als blaue Linie dargestellt. Im Gegensatz zu den Basisanforderungen wird der Kunde immer zufriedener, je mehr Leistungsanforderungen berücksichtigt werden. Dabei werden diese durch den Kunden vorgegeben.

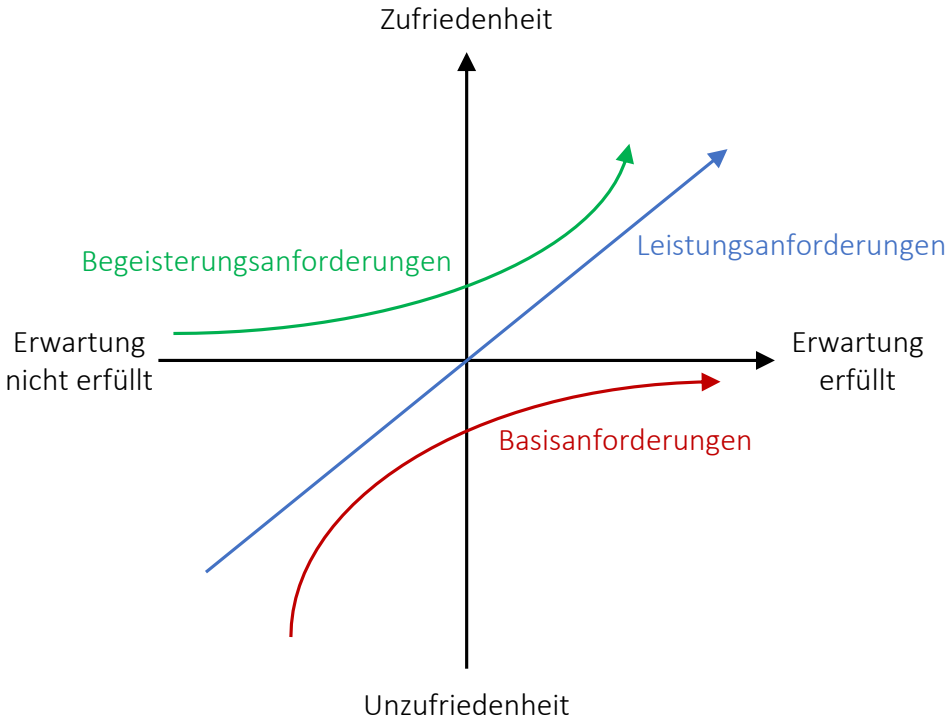


Abbildung A.1: Kano-Modell in Anlehnung an [7, S.85]

Die dritte Gruppe bilden die Begeisterungsanforderungen. Sie tragen stark zur Kundenzufriedenheit bei und sind dem Kunden noch gar nicht bewusst, dass er sie braucht. Durch die Begeisterungsanforderungen kann ein Produkt sich von etwaiger Konkurrenz abheben.

A.2 Fair Information Practice Principles

Nachfolgend sind die Fair Information Practice Principles (FIPPs) nach [10] tabellarisch aufgelistet. Die Prinzipen sind in den folgenden, teilweise bereits erwähnten Regulierungen zum Datenschutz verankert: DSGVO, den Australian Privacy Act's Information Privacy Principles, dem Gesetzentwurf zum Schutz personenbezogener Daten in Singapur sowie den indischen Information Technology Rules [440].

Tabelle A.1: Übersicht (A-M) der Fair Information Practice Principles (FIPPs) nach [10]

Prinzip	Erläuterung
Access and Amendment	Organisationen sollten Einzelpersonen einen angemessenen Zugang zu personenbezogenen Daten gewähren und ihnen die Möglichkeit geben, ihre Daten zu korrigieren oder zu ändern.
Accountability	Organisationen sollten für die Einhaltung dieser Grundsätze und der geltenden Datenschutzbestimmungen verantwortlich sein und die Einhaltung angemessen überwachen, prüfen und dokumentieren. Die Agenturen sollten auch die Rollen und Verantwortlichkeiten aller Mitarbeiter und Auftragnehmer in Bezug auf PII klar definieren und allen Mitarbeitern und Auftragnehmern, die Zugang zu PII haben, entsprechende Schulungen anbieten.
Authority	Organisationen sollten personenbezogene Daten nur dann erstellen, sammeln, verwenden, verarbeiten, speichern, aufbewahren, verbreiten oder offenlegen, wenn sie dazu befugt sind, und diese Befugnis in der entsprechenden Mitteilung angeben.
Individual Participation	Organisationen sollten den Einzelnen in den Prozess der Verwendung von PII einbeziehen und, soweit dies praktikabel ist, die Zustimmung des Einzelnen für die Erstellung, Erhebung, Verwendung, Verarbeitung, Speicherung, Pflege, Verbreitung oder Offenlegung von PII einholen. Organisationen sollten auch Verfahren zur Entgegennahme und Bearbeitung von Beschwerden und Anfragen zum Datenschutz einrichten.
Minimization	Organisationen sollten nur solche PII erstellen, sammeln, verwenden, verarbeiten, speichern, aufbewahren, verbreiten oder offenlegen, die unmittelbar relevant und notwendig sind, um einen gesetzlich genehmigten Zweck zu erfüllen, und sie sollten PII nur so lange aufbewahren, wie es für die Erfüllung des Zwecks erforderlich ist.

Tabelle A.2: Übersicht (P-T) der Fair Information Practice Principles (FIPPs) nach [10]

Prinzip	Erläuterung
Purpose Specification and Use Limitation	Organisationen sollten über den spezifischen Zweck, für den PII gesammelt werden, informieren und PII nur für einen Zweck verwenden, verarbeiten, speichern, aufbewahren, verbreiten oder offenlegen, der in der Mitteilung erläutert wird und mit dem Zweck, für den die PII gesammelt wurden, vereinbar ist oder der anderweitig gesetzlich genehmigt ist.
Quality and Integrity	Organisationen sollten personenbezogene Daten mit einer solchen Genauigkeit, Relevanz, Aktualität und Vollständigkeit erstellen, erheben, verwenden, verarbeiten, speichern, aufbewahren, verbreiten oder offenlegen, wie es vernünftigerweise erforderlich ist, um die Fairness gegenüber der Person zu gewährleisten.
Security	Organisationen sollten verwaltungstechnische, technische und physische Sicherheitsvorkehrungen zum Schutz von PII treffen, die dem Risiko und dem Ausmaß des Schadens entsprechen, der sich aus dem unbefugten Zugriff, der unbefugten Nutzung, Änderung, dem Verlust, der Zerstörung, der Verbreitung oder der Offenlegung von PII ergeben würde.
Transparency	Organisationen sollten ihre Informationspolitik und -praktiken in Bezug auf personenbezogene Daten transparent machen und klare und zugängliche Informationen über die Erstellung, Erhebung, Verwendung, Verarbeitung, Speicherung, Pflege, Verbreitung und Offenlegung von personenbezogenen Daten bereitstellen.

B

SLR und Kodierungsprozess - Konzept der Erklärbarkeit

B.1 Systematische Literaturrecherche

Bezugnehmend auf die in Abschnitt 4.1.1 beschriebene SLR.

B.1.1 Manuelle Suche

Folgende Literatur haben wir bei der manuellen Suche gefunden: [446, 447, 448, 162, 449, 299, 373, 406, 249, 163, 153, 450, 152, 451, 452, 453, 454, 455, 456, 457, 458, 207, 459, 460, 166, 409, 461, 52, 462, 463, 464, 465, 251, 466, 467, 468, 469, 470, 471, 472, 164, 473, 474, 475, 476, 477, 478, 479, 247, 246, 480, 481, 482, 161, 483, 484, 485, 486, 487, 488, 387, 489, 490, 491, 492, 493, 494, 415, 495, 496, 497, 498, 499, 500, 501, 259, 502, 503, 504, 505, 506, 167, 507, 397, 508, 509, 510, 511, 240, 261, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 372]

Tabelle B.1: Quellen und Auswahl der Publikationen der manuellen Suche

Quelle	Gesamt[1]	Initialauswahl	Finale Auswahl
International Requirements Engineering Conference (RE)	1312	15	4
Symposium on the Foundations of Software Engineering (FSE)	667	8	2
Information and Software Technology (IST)[2]	2668	23	0
Intelligent User Interfaces (IUI)	1158	52	18
Journal of Systems and Software (JSS)[3]	4121	8	0
Transaction on Software Engineering (TSE)	2910	23	0
Conference on Information and Knowledge Management (CIKM)	2789	8	2
International Working Conference on Requirement Engineering: Foundation for Software Quality (REFSQ)	328	4	0
Transactions on Software Engineering and Methodology (TOSEM)	615	4	1
Conference on Recommender Systems (RecSys)	521	15	6
Requirements Engineering Journal (REJ)	455	9	2
RE Workshops	21	4	1
REFSQ Workshops[4]	162	5	1
Minds and Machines	724	30	17
Big Data & Society	284	16	6
International Joint Conference on Artificial Intelligence - Workshop on eXplainable Artificial Intelligence[5]	63	41	34
Philosophy and Technology[6]	259	10	9
Ethics and Information Technology	538	1	1

[1] Gesamtanzahl an überprüften Arbeiten je Quelle.

[2] Kein Zugriff auf 20 von 2668 Publikationen.

[3] Kein Zugriff auf 7 von 4121 Publikationen.

[4] Nur die Proceedings der Jahre 2000, 2001, 2006-2008, 2010, 2011 und 2015-2019 waren zugänglich.

[5] Proceedings von 2017-2019 wurden berücksichtigt

[6] Nur Ausgaben ab 2011 waren zugänglich.

B.1.2 Snowballing

Folgende Literatur wurde mit Hilfe des Snowballings identifiziert: [525, 526, 527, 528, 529, 530, 531, 68, 252, 248, 532, 533, 534, 535, 536, 417, 253, 537, 538, 238, 413, 539, 540, 541, 414, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 257, 218, 555, 556, 557, 558, 559, 560, 407, 561, 562, 403, 563, 564, 565, 566, 567, 568, 569, 570, 155, 571, 572, 573, 574, 575, 576, 577, 412, 578, 579, 580, 581, 168, 582, 239, 583, 584, 585, 250, 586, 237, 587, 258, 214, 260, 588, 151, 589, 150, 590, 241, 591, 592, 593, 594, 595, 596, 597, 598, 217, 599, 254, 600, 601, 602, 603, 604, 216, 605, 606, 343, 607, 608, 609, 610, 611, 612, 418, 264, 613, 614, 615, 616, 617, 618, 244, 439]

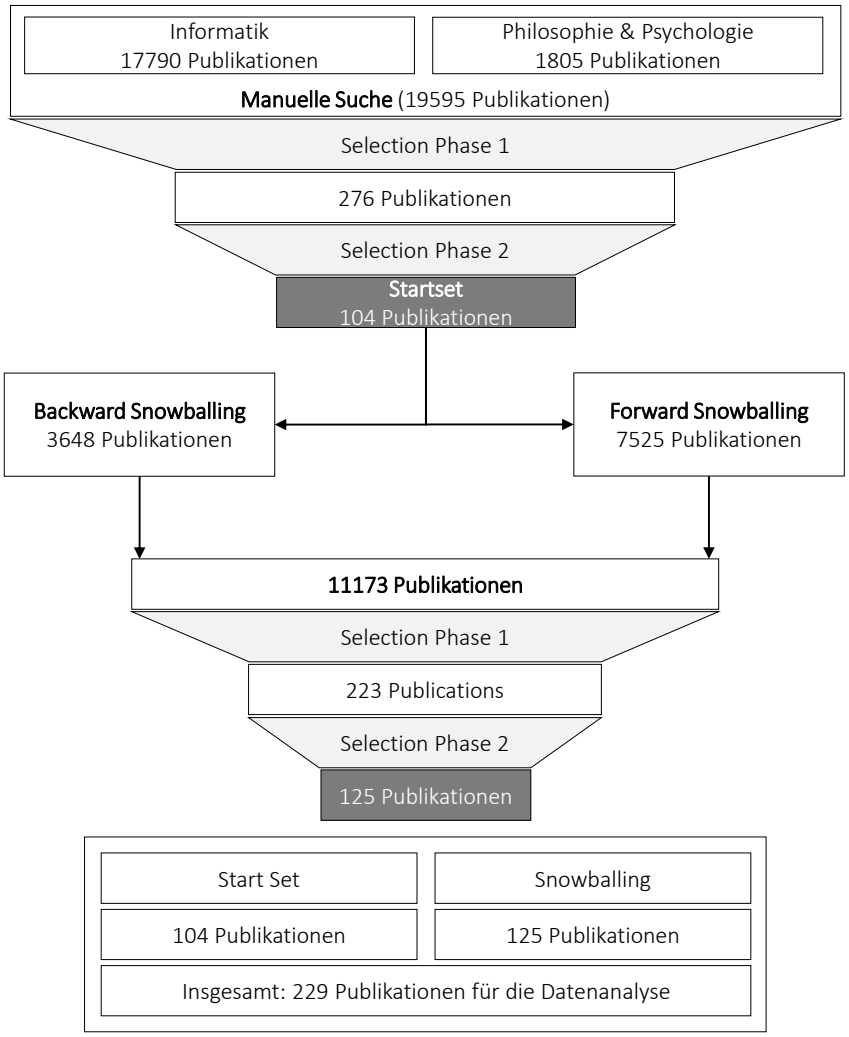


Abbildung B.1: Grafische Übersicht der SLR in Anlehnung an [3]

B.1.3 Auswahlprozess der Publikationen

Einschluss- und Ausschlusskriterien Wir haben Publikationen ausgewählt, die die folgenden Einschlusskriterien erfüllten:

IC_1 Adressieren mind. eine unserer Forschungsfragen

IC_2 Sind im Zeitraum 01/1984 – 03/2020 veröffentlicht worden

IC_3 Peer-reviewed Journal-, Konferenz- oder Workshop-Beiträge

Wurde mind. eines der folgenden Ausschlusskriterien nicht erfüllt, haben wir die Publikation nicht berücksichtigt:

EC_1 Nicht in englischer Sprache verfasst

EC_2 Tutorials, Anträge oder andere non-peer reviewed Publikation

EC_3 Arbeiten, die rein algorithmische Techniken erforschen oder vorschlagen, ohne weitere Diskussion über den theoretischen Hintergrund der Erklärbarkeit

Wir haben 1984 als Startdatum gewählt, da in diesem Jahr die erste größere Arbeit über Erklärbarkeit veröffentlicht wurde [212]. Wir sind uns durchaus der Tatsache bewusst, dass 36 Jahre eine lange Zeitspanne sind, wollten aber gezielt durch die Wahl dieser Zeitspanne einen möglichst breiten Überblick über das Thema gewinnen. Damit eine Publikation ausgewählt wurde, mussten alle Einschlusskriterien erfüllt sein. War mindestens eines der Ausschlusskriterien erfüllt, wurde diese Publikation abgelehnt. Formal ausgedrückt:

$$(IC_1 \wedge IC_2 \wedge IC_3) \wedge \neg(EC_1 \vee EC_2 \vee EC_3)$$

B.1.4 Codes und Details zu den Workshops

Für Details zu den von uns identifizierten Codes aus [4] und den Details zu den von uns durchgeführten Workshops verweise ich auf [619], wo alle Daten einsehbar sind.

B.2 Qualitätsmerkmale Erklärbarkeit und Privatsphäre

Nachfolgend sind die Qualitätsmerkmale der Datenanalyse aus meiner in Abschnitt 4.1.2.1 beschriebenen LR zu finden, die im Beziehungsmodell Erklärbarkeit und Privatsphäre (siehe Abschnitt 4.3) Anwendung finden.

Tabelle B.2: Überblick der Qualitätsmerkmale aus der LR mit Quellenabgabe (A-G)

Qualitätsaspekt	**Quelle(n)**
Accessibility	[48, 282, 122, 174, 270, 45, 297, 118, 620, 281, 439]
Accountability	[274, 271, 122, 265, 270, 280, 41, 273, 277]
Actionable	[184, 187, 140, 174, 114, 620, 272, 439]
Adaptability	[187, 266, 272, 439]
Compliance	[184, 274, 271, 122, 278, 174, 265, 279, 293, 280, 41, 276, 273, 277, 439]
Comprehensibility	[175, 299, 268, 187, 186, 140, 269, 164, 122, 291, 174, 265, 287, 270, 293, 114, 45, 620, 266, 281]
Contextuality	[189, 187, 291, 174, 192, 280, 266, 439]
Correctness	[140, 122, 291, 297, 275]
Customer Loyalty	[184, 271, 288, 281]
Customer Trust	[184, 291, 278, 288, 281]
Decision Making	[189, 184, 299, 283, 268, 394, 187, 48, 282, 186, 140, 164, 267, 291, 262, 174, 288, 192, 279, 49, 114, 45, 193, 273, 621, 620, 266, 271, 276, 439]
Educate	[140, 267, 114, 45, 272]
Empower	[189, 174, 290, 114, 45, 297, 266]
Ethics	[263, 184, 274, 271, 280, 41, 273, 281]
Experienced Effectiveness	[288, 192, 114, 140, 279]
Experienced Usefulness	[189, 184, 282, 186, 140, 164, 291, 262, 288, 192, 279, 49, 109]
Experienced Value	[184, 271, 282, 186, 140, 111, 164, 267, 291, 262, 288, 192, 49, 118, 621, 266, 272, 281, 109]
Fairness	[263, 184, 274, 283, 271, 48, 282, 270, 279, 41, 276, 277, 275]
Guidance	[186, 140, 122, 262, 174, 270, 41]

Tabelle B.3: Überblick der Qualitätsmerkmale aus der LR mit Quellenabgabe (I-Z)

Qualitätsaspekt	Quelle(n)
Information Granularity	[189, 184, 187, 164, 291, 439]
Informativeness	[175, 184, 299, 283, 282, 186, 291, 262, 174, 265, 270, 41, 45, 266, 281, 109]
Lawful	[274, 270, 41, 276, 273, 277, 275]
Privacy Awareness	[189, 111, 287, 114, 118, 187, 186, 140, 267, 262, 265, 41, 266, 272, 281]
Transparency	[189, 184, 282, 269, 164, 122, 262, 278, 265, 288, 192, 287, 270, 279, 290, 49, 41, 45, 276, 273, 281, 277, 109, 275, 439]
Trust	[263, 175, 299, 271, 48, 282, 186, 140, 111, 164, 291, 262, 278, 174, 288, 192, 287, 270, 280, 49, 41, 118, 621, 266, 281, 277, 275]
Trustworthiness	[263, 184, 299, 282, 111, 164, 288, 287, 49, 273, 298, 266, 272]
Understandability	[263, 175, 184, 289, 268, 48, 282, 186, 140, 269, 122, 291, 262, 174, 265, 288, 287, 279, 49, 45, 297, 193, 273, 298, 266, 281, 109, 439]
Usability	[187, 174, 192, 49, 114, 45, 620, 263, 175, 184, 140, 267, 291, 262, 41, 118, 193, 109, 48, 282, 186, 279, 280, 297, 273, 281, 439]
User Acceptance	[175, 282, 291, 265, 288, 114, 41, 621, 266, 281, 109, 278, 293, 273]
User Control	[189, 184, 299, 283, 289, 268, 187, 282, 186, 111, 291, 174, 288, 192, 287, 270, 279, 290, 280, 114, 276, 193, 273, 621, 298, 620, 266, 281, 109, 439]
User Efficiency	[268, 187, 282, 186, 278, 192, 280, 272]
User Experience	[394, 282, 186, 140, 291, 262, 279, 49, 114, 41, 621, 272, 281, 109]
User Satisfaction	[186, 140, 164, 287, 49, 114]
Verifiability	nur von den Experten der Fokusgruppe genannt

B.2.1 Modellvalidierung mit der Fokusgruppe

B.2.1.1 Vorbereitungsaufgaben

Aufgabe 1. Zu Beginn bitte ich Euch – sofern das nicht ohnehin schon der Fall ist – sich nochmal mit den Dimensionen des Conceptual Models aus dem „Exploring Explainability" Paper [4] vertraut zu machen.

Aufgabe 2. Schaut Euch bitte die mitgesendete Liste der Qualitätsmerkmale einmal an. Macht Euch bitte Gedanken dazu, ob Eurer Meinung nach hier noch Qualitätsmerkmale fehlen oder ob Eurer Meinung nach einige vielleicht gar nicht in den Beziehungskontext von Erklärbarkeit und Privacy gehören. Eine kleine Hilfestellung, um sich ein wenig in die Thematik hinein zu versetzen, sollen die beiden Szenarien (Abschnitt B.2.1.2) hier leisten.

B.2.1.2 Szenarien der Vorbereitungsaufgabe

Alice im Smart Home Alice wohnt in einem Smart Home, leidet unter einer chronischen Nierenerkrankung und hat zu Hause ein Hämodialysegerät zur Therapie. Nun möchte der örtliche Stromversorger gerne Verbrauchsdaten der Bewohner sammeln, um möglichst ökologisch und ökonomisch den Strom in der Stadt zu verteilen. Darüber hinaus kann er den Verbrauchern so auch detaillierte Energiespartips geben, damit diese Kosten und Verbrauch reduzieren können. Grundsätzlich ist Alice an den Verbrauchertips des Stromversorgers interessiert und ist bereit feingranulare Daten über ihren Stromverbrauch mit dem Versorger zu teilen. Allerdings möchte Alice das nicht, wenn sie ihr Hämodialysegerät verwendet, da sie ihrem Stromversorger keine Auskünfte über ihren Gesundheitszustand zukommen lassen möchte.

Sightseeing App Sie haben über einem Bekannten von einer neuen App gehört, mit der Sie Sightseeing-Touren oder Tagesausflüge für Städte auf der ganzen Welt planen können. Sie beschließen, diese App für Ihren bevorstehenden Städtetrip auf Ihr Smartphone herunterzuladen. Beim ersten Start der App fragt diese jedoch, ob Ihr Standort verwendet werden darf und fragt auch nach Ihrem Geburtsdatum. Sie sind sich über den Grund nicht sicher, da die App Ihnen keine weiteren Informationen über die Verwendung Ihrer Daten gibt.

B.2.1.3 Gruppenaufgaben

Aufgabe 1. Jeder von Euch wählt sich bitte ein Qualitätsmerkmal aus der alphabetisch sortierten Liste aus, ordnet es einer Dimension zu und begründet seine Entscheidung. Anschließend diskutieren alle gemeinsam die Entscheidung mit dem Ziel, einen Konsens zu finden.

Aufgabe 2. Schaut Euch bitte folgendes von mir erstelltes Modell in Ruhe an. Vergleicht es anschließend mit dem gerade erstellten Modell, identifiziert Unterschiede und diskutiert diese anschließend mit der Gruppe.

C

Ergänzendes Material zum Privacy Policy eXplainer

C.1 Icons der Kategorien von PriX

Abbildung C.1: Zuordnung von Privacy Icons und Kategorien

C.2 Erweiterte Benutzeroberfläche von PriX

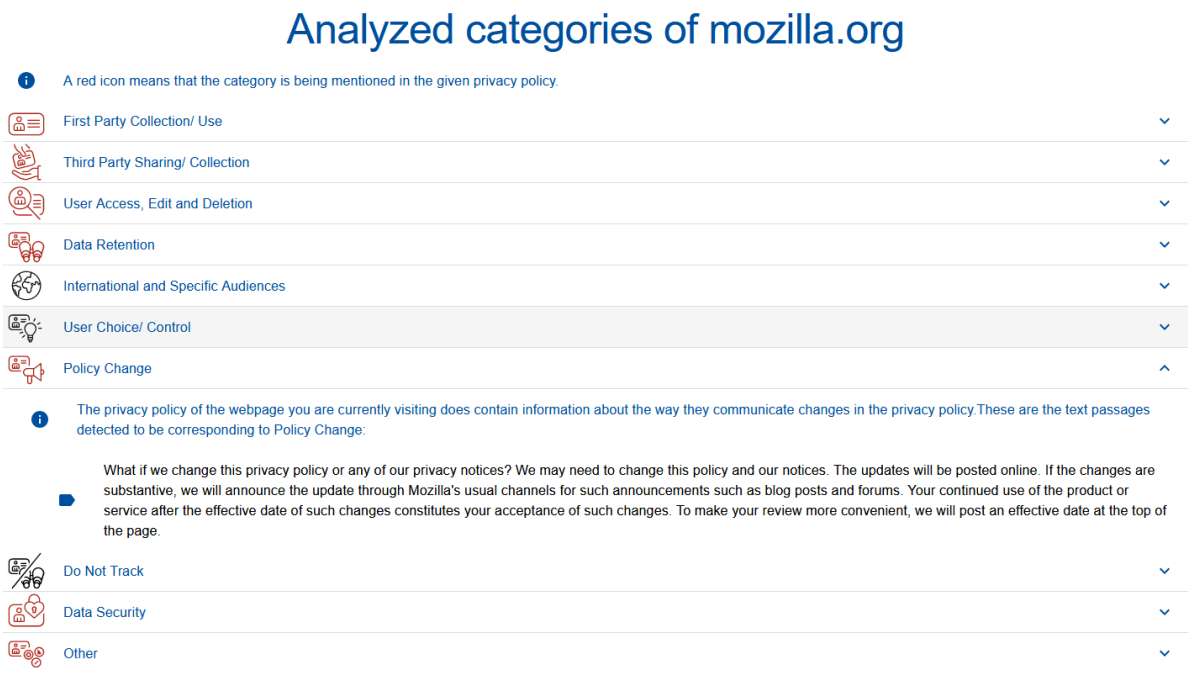


Abbildung C.2: Details Benutzeroberfläche von PriX

C.3 Kategorien des OPP-115 Korpus

First Party Collection (FPC) Informationen über Daten, die vom Betreiber der Internetseite gesammelt werden und wie diese genutzt werden.	**Third Party Sharing/ Collection (TPS)** Informationen darüber, ob Daten mit Dritten getauscht werden und in welchem Umfang dies geschieht.
Data Security (DS) Informationen darüber, ob und wie vom Betreiber gespeicherte Daten geschützt werden.	**User Access, Edit and Deletion (UAED)** Informationen darüber, wie der Nutzer auf seine Daten zugreifen und diese bearbeiten oder löschen kann.
User Choice/ Control (UCC) Informationen darüber, ob und wie weit der Nutzer selber entscheiden kann, welche Daten über ihn gespeichert werden dürfen.	**Data Retention (DR)** Informationen darüber, wie und wie lange Daten seitens des Betreibers der Internetseite gespeichert werden.
Policy Change (PC) Informationen darüber, wie Veränderungen in der Datenschutzerklärung veröffentlicht werden.	**Do Not Track (DNT)** Informationen darüber, wie mit Do Not Track-Signalen umgegangen wird.
International and Specific Audiences (ISA) Informationen für Besucher aus dem Ausland und spezifische Personengruppen.	**Other (O)** Andere Informationen, wie beispielsweise Möglichkeiten zur Kontaktaufnahme mit dem Betreiber der Internetseite.

Tabelle C.1: Die analysierten Kategorien

D

Survey-Material Privacy Explanations

D.1 Survey Instrument

Umfrage zu Erklärungen der Privatsphäre in Software-Systemen

Herzlich Willkommen!

Vielen Dank, dass Sie sich die Zeit genommen haben, an unserer Umfrage teilzunehmen. Mit Hilfe dieser Umfrage soll die aktuelle Wahrnehmung der Endbenutzer in Bezug auf die Auswirkung von Erklärungen zu Aspekten der Privatsphäre bei Software-Anwendungen untersucht werden. Das Ziel ist es, zu verstehen, wie Ihre aktuelle Wahrnehmung in Bezug auf Ihre Privatsphäre bei Software-Anwendungen, die Sie in Ihrem täglichen Leben verwenden, ist und ob Sie daran interessiert sind, Erklärungen zu Aspekten der Privatsphäre, sogenannte *Privacy Explanations*, während der Verwendung einer Software-Anwendung zu erhalten.

Bitte beantworten Sie alle Fragen ehrlich, denn authentische Antworten sind entscheidend für den Erfolg dieser Untersuchung. Es gibt keine richtige oder falsche Antwort, Sie können also ohne Bedenken Ihre wahre Meinung äußern.

Diese Umfrage ist vollständig anonym. Die Antworten auf die Fragen lassen keine Rückschlüsse auf Ihre Person zu.

Durch die Teilnahme an der Umfrage bestätigen Sie, dass Sie volljährig sind (mindestens 18 Jahre alt). Diese Umfrage umfasst 30 Fragen und die Beantwortung der Fragen dauert im Durchschnitt 25 Minuten.

A. Demografische Daten

[A1] In welchem Land haben Sie Ihren Wohnsitz?
<*Länderliste als Auswahloption*>

[A2] Mit welcher Geschlechtsidentität identifizieren Sie sich am meisten?

- Weiblich
- Männlich
- Transgender Weiblich
- Transgender Männlich
- Gender Variant/nicht übereinstimmend
- Möchte nicht antworten
- Nicht aufgeführt

[A3] Welche der folgenden Angaben beschreibt Ihre derzeitige berufliche Tätigkeit am besten?

- Bau, Architektur, Vermessung und Gebäudetechnik
- Gesundheit, Soziales, Lehre und Erziehung
- Im Ruhestand
- Kaufmännische Dienstleistungen, Warenhandel, Vertrieb, Hotel und Tourismus
- Land-, Forst- und Tierwirtschaft und Gartenbau
- Militär
- Naturwissenschaft, Geografie und Informatik
- Rohstoffgewinnung, Produktion und Fertigung
- Sprach-, Literatur-, Geistes-, Gesellschafts- und Wirtschaftswissenschaften, Medien, Kunst, Kultur und Gestaltung
- Student
- Unternehmensorganisation, Buchhaltung, Recht und Verwaltung
- Verkehr, Logistik, Schutz und Sicherheit
- Sonstiges (bitte angeben)

[A4] In welchem Jahr wurden Sie geboren?

Geben Sie das Datum ein (z.B. 1991): _____

[A5a] Was ist Ihr Familienstand?

- Single, nie verheiratet
- Verheiratet oder in einer Lebenspartnerschaft
- Verwitwet
- Geschieden
- Getrennt
- Sonstiges (bitte angeben)

[A5b] Was ist Ihr Familienstand?

Privacy Explanation: *Mit dieser Frage soll untersucht werden, ob der Beziehungsstatus einen Einfluss auf die Wahrnehmung der Privatsphäre hat.*

- Single, nie verheiratet
- Verheiratet oder in einer Lebenspartnerschaft
- Verwitwet
- Geschieden
- Getrennt
- Sonstiges (bitte angeben)

[A5c] Was ist Ihr Familienstand?

Zeige Privacy Explanation

- Single, nie verheiratet
- Verheiratet oder in einer Lebenspartnerschaft
- Verwitwet
- Geschieden
- Getrennt
- Sonstiges (bitte angeben)

[A6] Haben Sie Kinder?

abhängig von Antwort der Frage A5

- Ja
- Nein

B. Software Kenntnisse

[B1] Bitte markieren Sie alle Punkte, die auf Sie zutreffen:

- Ich habe nicht so viel Erfahrung, aber ich kann meine E-Mails abrufen und einfache Aufgaben mit Textverarbeitungsprogrammen erledigen.
- Ich bin versiert im Umgang mit Textverarbeitungs-Software (LibreOffice, Microsoft Office, OpenOffice, etc.). Ich bin sehr sicher im Umgang mit diesen Programmen und habe viel Erfahrung damit.
- Ich kann Begriffe wie Hardware und Software unterscheiden.
- Ich bin sehr sicher im Umgang mit Computern. Ich kann mich schnell in neue Programme einarbeiten.
- Ich bin in der Lage, komplexe Aufgaben durchzuführen, wie das Erstellen von funktionalen Tabellenkalkulationen und Organisieren und Analysieren großer Datenmengen.
- Ich verstehe die grundlegenden Konzepte einer Web-Seite: URL, *header*, Suchfeld, Hyperlinks, *footer*.
- Ich verstehe Netzwerke Konzepte wie IP-Adresse, VPN, Router, LANs.
- Ich verstehe Konzepte von Betriebssystemen, wie CPU-Nutzung, Speicherverbrauch und Festplattenkapazität.
- Ich bin der Meinung, dass ich über grundlegende oder fortgeschrittene Programmierkenntnisse verfüge.
- Ich bin Software-/Hardware-Entwickler.

[B2] Welche der folgenden Geräte nutzen Sie im Alltag?
(Mehr als eine Auswahl erlaubt)

- Laptop/Notebook/Netbook
- Smartphone
- Tablet
- Desktop Computer
- Smartwatch
- Sonstiges (bitte angeben)

[B3] Verwenden Sie an einem typischen Tag Software-Anwendungen eher aus beruflichen oder privaten Gründen?

- Eher aus beruflichen Gründen
- Etwas mehr aus beruflichen Gründen
- Etwa zu gleichen Teilen beruflichen und privaten Gründen
- Etwas mehr aus privaten Gründen
- Eher aus privaten Gründen

[B4] Welche Kategorie von Software/Apps nutzen Sie an einem typischen Tag am häufigsten auf Ihren digitalen Geräten?

(Mehr als eine Auswahl erlaubt)

	1 - Nie	2 - Fast nie	3 - Gelegentlich	4 - Fast immer	5 - Sehr oft
E-Mail Client (Microsoft Outlook, Thunderbird)	○	○	○	○	○
Unterhaltung (Musik, Videos, etc.)	○	○	○	○	○
Spiele (Browser, Konsole, PC, etc.)	○	○	○	○	○
Gesundheit/Fitness Apps	○	○	○	○	○
IDE (Integerated Development Environment)	○	○	○	○	○
Management Systeme (Buchhaltung, Vertrieb, etc.)	○	○	○	○	○
Mobilität (Navigation, Öffentliche Verkehrsmittel, etc.)	○	○	○	○	○
Nachrichten	○	○	○	○	○
Organizers (Kalender, ToDo Listen, etc.)	○	○	○	○	○
Outdoor Apps/Tour Guides	○	○	○	○	○
Elektronische Nachschlagewerke (Suchmaschinen/Enzyklopädien)	○	○	○	○	○
Shopping Portale (Apps und Websites)	○	○	○	○	○
Social Networking/Messengers(Twitter, Facebook, etc.)	○	○	○	○	○
Textverarbeitung	○	○	○	○	○
Browser (Chrome, Edge, Firefox, Internet Explorer, Safari, etc.)	○	○	○	○	○

C. Privatsphäre im Allgemeinen

[C1] Lesen Sie jeden Punkt sorgfältig durch und wählen Sie dann für jede der folgenden Aussagen, wie stark Sie zustimmen oder nicht zustimmen.

	1 - Stimme überhaupt nicht zu	2 - Stimme nicht zu	3 - Stimme eher nicht zu	4 - Stimme weder zu noch nicht zu	5 - Stimme eher zu	6 - Stimme zu	7 - Stimme völlig zu
Die Verbraucher haben jegliche Kontrolle darüber verloren, wie persönliche Informationen von Unternehmen gesammelt und verwendet werden.	○	○	○	○	○	○	○
Die meisten Unternehmen gehen mit den persönlichen Daten, die sie über Verbraucher sammeln, ordnungsgemäß und vertraulich um.	○	○	○	○	○	○	○
Bestehende Gesetze und Organisationspraktiken bieten heute ein angemessenes Maß an Schutz für die Privatsphäre der Verbraucher.	○	○	○	○	○	○	○

D. Eigene Wahrnehmung der Privatsphäre

[D1] Lesen Sie jeden Punkt sorgfältig durch und wählen Sie dann für jede der folgenden Aussagen, wie stark Sie zustimmen oder nicht zustimmen.

	1 - Stimme überhaupt nicht zu	2 - Stimme nicht zu	3 - Stimme eher nicht zu	4 - Stimme weder zu noch nicht zu	5 - Stimme eher zu	6 - Stimme zu	7 - Stimme völlig zu
Ich fühle mich sicher, bei der Weitergabe meiner persönlichen Daten an Online-Dienste (wie Online-Shops) und/oder Apps.	○	○	○	○	○	○	○
Die Bereitstellung von persönlichen Informationen an Online-Dienste oder Apps verursacht zu viele Bedenken.	○	○	○	○	○	○	○
Im Allgemeinen vertraue ich Online-Unternehmen im Umgang mit meinen persönlichen Daten, z.B. meiner Kaufhistorie.	○	○	○	○	○	○	○

[D2] Wie besorgt sind Sie über die Bedrohungen Ihrer persönlichen Privatsphäre im Internet heutzutage?

Zum Beispiel:

„Dritte könnten Zugriff auf meine persönlichen Daten erhalten“ oder „Mein Konsumverhalten könnte Dritten bekannt werden“

- ☐ 1 - Überhaupt nicht besorgt
- ☐ 2 - Kaum besorgt
- ☐ 3 - Ein wenig besorgt
- ☐ 4 - Mäßig besorgt
- ☐ 5 - Stark besorgt

[D3] Können Sie uns EINE Bedrohung nennen, über die Sie besonders besorgt sind?

<*Freitextantwort*>

E. Privatsphäre und Software

[E1] Wie oft achten Sie bei der Installation von Apps auf Ihrem Smartphone auf *„erforderliche Berechtigungen“* (z.B. Zugriff auf Mikrofon, Kamera usw.)?

- ☐ 1 - Nie
- ☐ 2 - Manchmal
- ☐ 3 - Immer

[E2] Wie schnell klicken Sie bei der Installation von Software die Schaltfläche *„Zustimmen“*?

- ☐ 1 - Sofort
- ☐ 2 - Innerhalb einer Minute
- ☐ 3 - Dauert länger als eine Minute

[E3] Achten Sie generell darauf, ob eine Website, die Sie besuchen, eine Datenschutzerklärungen hat oder nicht?

- ☐ 1 - Niemals
- ☐ 2 - So gut wie nie
- ☐ 3 - Selten
- ☐ 4 - Manchmal
- ☐ 5 - Häufig
- ☐ 6 - Sehr häufig
- ☐ 7 - Immer

[E4] Wie oft lesen Sie die Datenschutzerklärungen von Websites?

- ☐ 1 - Niemals
- ☐ 2 - So gut wie nie
- ☐ 3 - Selten
- ☐ 4 - Manchmal
- ☐ 5 - Häufig
- ☐ 6 - Sehr häufig
- ☐ 7 - Immer

F. Probleme bei der Nutzung von Software im Kontext der Privatsphäre

[F1] Wie oft haben Sie ein ungutes Gefühl in Bezug auf den Datenschutz, wenn Sie Software verwenden oder Websites besuchen, die mit diesen Kategorien zusammenhängen?

	1 - Nie	2 - Sehr selten	3 - Selten	4 - Gelegentlich	5 - Häufig	6 - Sehr häufig
Unterhaltung (Musik, Videos, etc)	○	○	○	○	○	○
Spiele	○	○	○	○	○	○
Gesundheit/Fitness Apps	○	○	○	○	○	○
Mobilität (Navigation, Öffentliche Verkehrsmittel, etc.)	○	○	○	○	○	○
Organizers (Kalender, ToDo Listen, etc.)	○	○	○	○	○	○
Outdoor Apps/Tour Guides	○	○	○	○	○	○
Elektronische Nachschlagewerke	○	○	○	○	○	○
Shopping Portale (Apps und Websites)	○	○	○	○	○	○
Social Networking/Messengers	○	○	○	○	○	○

[D3] Können Sie uns EINE Situation nennen, in der Sie sich bei der Verwendung von Software in Bezug auf die Privatsphäre unwohl gefühlt haben? Bitte beginnen Sie damit, zu sagen, um welche Art von Software (Kategorie oder Name) es sich handelte.
Seien Sie so spezifisch wie möglich und fassen Sie sich kurz.
<Freitextantwort>

G. Bedarf an Erklärungen zur Privatsphäre (Privacy Explanations)

[G1] Hypothetische Situation: Sie haben über einem Bekannten von einer neuen App gehört, mit der Sie Sightseeing-Touren oder Tagesausflüge für Städte auf der ganzen Welt planen können. Sie beschließen, diese App für Ihren bevorstehenden Städtetrip auf Ihr Smartphone herunterzuladen. Beim ersten Start der App fragt diese jedoch, ob Ihr Standort verwendet werden darf und fragt auch nach Ihrem Geburtsdatum. Sie sind sich über den Grund nicht sicher, da die App Ihnen keine weiteren Informationen über die Verwendung Ihrer Daten gibt. Wie sehr sind Sie daran interessiert, eine Erklärung in Bezug auf Ihre Privatsphäre (Privacy Explanation) zu erhalten?

- ☐ 1 - Gar nicht interessiert
- ☐ 2 - Geringfügig interessiert
- ☐ 3 - Indifferent
- ☐ 4 - Mäßig interessiert
- ☐ 5 - Sehr interessiert
- ☐ 6 - Extrem interessiert

[G2] Betrachten Sie die gerade gezeigte hypothetische Situation. Die Software gibt Ihnen nun jeweils Erklärungen.

1. Die Erklärung in Bezug auf die Nutzungs des Standortes: *Um Ihnen Touren und Empfehlungen in Ihrer Nähe anzeigen zu können, benötigen wir Zugriff auf Ihren Standort.*

2. Die Erklärung bzgl. der Verwendung des Geburtsdatums lautet: *Anhand Ihres Geburtsdatums können wir Ihnen Empfehlungen zeigen, was anderen Nutzern Ihres Alters gefallen hat.*

Wie nützlich finden Sie diese Art von Privatsphäreerklärungen?

1. Die Standortabfrage betreffend:

- ☐ 1 - Absolut nutzlos
- ☐ 2 - Ziemlich nutzlos
- ☐ 3 - Geringfügig nutzlos
- ☐ 4 - Weder nützlich noch nutzlos
- ☐ 5 - Geringfügig nützlich
- ☐ 6 - Ziemlich nützlich
- ☐ 7 - Absolut nützlich

2. In Bezug auf das Geburtsdatum:

- ☐ 1 - Absolut nutzlos
- ☐ 2 - Ziemlich nutzlos
- ☐ 3 - Geringfügig nutzlos
- ☐ 4 - Weder nützlich noch nutzlos
- ☐ 5 - Geringfügig nützlich
- ☐ 6 - Ziemlich nützlich
- ☐ 7 - Absolut nützlich

[G3] Betrachten Sie die zuvor gezeigte hypothetische Situation mit den gegebenen Privacy Explanations. Fühlen Sie sich mit diesen Erklärungen jetzt wohler, da Sie wissen, wofür die App Ihre Daten benötigt?

- ☐ Ja, ich fühle mich wohler.
- ☐ Es hat mich vorher schon nicht beunruhigt.
- ☐ Nein, weil: <*Freitextantwort*>

[G4] Stimmen Sie zu, dass Privacy Explanations ein möglicher Faktor sein können, um das Maß an Vertrauen in ein Software-System zu erhöhen?

- ☐ 1 - Stimme überhaupt nicht zu
- ☐ 2 - Stimme nicht zu
- ☐ 3 - Stimme eher nicht zu
- ☐ 4 - Stimme weder zu noch nicht zu
- ☐ 5 - Stimme eher zu
- ☐ 6 - Stimme zu
- ☐ 7 - Stimme völlig zu

[G5] Sind Sie generell an Privacy Explanations (z.B. zur Verwendung Ihrer Daten) interessiert?

- ☐ 1 - Gar nicht interessiert
- ☐ 2 - Geringfügig interessiert
- ☐ 3 - Indifferent
- ☐ 4 - Mäßig interessiert
- ☐ 5 - Sehr interessiert
- ☐ 6 - Extrem interessiert

[G6] Was erwarten Sie von einer Privacy Explanations bzw. was müsste diese für Sie enthalten? Seien Sie so konkret und anschaulich wie möglich.
<*Freitextantwort*>

H. Meinung/Ansicht zu Privacy Explanations

[H1] Was ist Ihrer Meinung nach der Nutzen von Privacy Explanations? Nennen Sie bitte bis zu drei mögliche Punkte, die ihnen dazu einfallen

- ☐ 1. <*Freitextantwort*>
- ☐ 2. <*Freitextantwort*>
- ☐ 3. <*Freitextantwort*>

[H2] Wann sollte eine Privacy Explanation angezeigt werden?

- ☐ Jedes Mal, wenn ich sie anfordere
- ☐ Automatisch, wenn etwas in Bezug auf meine Privatsphäre passiert (z.B. Nutzung von persönlichen Daten)
- ☐ Niemals
- ☐ Sonstiges: <*Freitextantwort*>

E

Konzept und technische Umsetzung von Privacy Explanations

E.1 Literaturrecherche

Bei unserer LR orientierten wir uns an den Richtlinien für SLRs von Kitchenham [224], den Richtlinien für das vorwärts und rückwärts Snowballing von Wohlin [225] sowie dem auf GT gründendem Ansatz von Wolfswinkel et al. [229]. Unsere LR bestand aus einer manuellen Inspektion, Snowballing und einer Datenbanksuche. Abbildung 7.2 veranschaulicht diesen Prozess. Nachfolgend sind die einzelnen Schritte und das Vorgehen ausführlich erläutert.

In allen drei LR-Phasen (manuelle Inspektion, Snowballing und Datenbanksuche) haben wir Publikationen auf Basis von Titel, Abstract und Keywords vorselektiert. Ließen diese Daten noch keinen hinreichenden Schluss zu, haben wir zudem die Conclusion miteinbezogen. Für die finale Auswahl, ob die Arbeit von Relevanz ist, wurde auch der gesamte Inhalt des Papers berücksichtigt. Sowohl bei der Vorselektion als auch der finalen Selektion bezogen wir unsere Einschluss- und Ausschlusskriterien mit ein (siehe Abschnitt E.1.4).

E.1.1 Manuelle Inspektion

Als Ausgangspunkt für unsere LR haben wir unsere bereits durchgeführte SLR [4] genutzt. Es ist unserem Kenntnisstand nach die aktuellste und vollständigste Literaturstudie zum Konzept der Erklärbarkeit im Bereich des SE bzw. RE. Wir prüften die von uns identifizierten Paper auf Relevanz bezüglich unseren Einschluss- und Ausschlusskriterien (Abschnitt E.1.4). Die final als relevant selektierten Publikationen bildeten unser Startset. Folgende 42 Publikationen haben wir bei der manuellen Inspektion gefunden: [622, 612, 214, 245, 555, 515, 216, 4, 167, 165, 254, 565, 240, 419, 217, 238, 168, 52, 154, 623, 243, 161, 241, 247, 155, 252, 164, 408, 520, 481, 404, 218, 624, 625, 292, 409, 409, 471, 410, 397, 516, 588]

E.1.2 Snowballing

Wir haben Vorwärts- und Rückwärts-Snowballing für die Publikationen aus unserem Startset durchgeführt. Bereits zu Beginn der zweiten Iteration haben wir eine Sättigung in unseren Daten festgestellt, so dass wir gemäß dem GT-Ansatz daraufhin stoppten. Folgende 45 Publikationen haben wir während des Snowballings gefunden: [626, 627, 628, 237, 629, 562, 398, 630, 631, 632, 633, 259, 634, 635, 616, 636, 248, 553, 637, 638, 639, 640, 403, 484, 641, 475, 642, 343, 643, 644, 564, 449, 520, 572, 645, 162, 184, 257, 258, 169, 646, 548, 647, 648, 421]

E.1.3 Datenbanksuche

Die Mehrheit der Publikationen aus der manuellen Inspektion und des Snowballings beschäftigen sich mit dem Konzept der Erklärbarkeit also solches und nicht immer im direkten Bezug zur Privatsphäre oder Vertrauen von Endbenutzern. Daher haben wir uns entschieden, zusätzlich eine Datenbanksuche durchzuführen mit ausschließlichem Bezug von Erklärbarkeit und Privatsphäre. Für die Datenbanksuche haben wir Google Scholar verwendet. Wie in Abschnitt 4.1.2.1 bereits diskutiert ist Google Scholar für Datenbanksuchen bei (systematischen) Literaturrecherchen geeignet [234]. Wir haben insgesamt drei Datenbanksuchen mit jeweils drei unterschiedlichen Such-Strings durchgeführt. Der Grund dafür ist zum einen, das ein zu generalisierter Such-String auch eine sehr hohe Anzahl an möglicherweise nicht relevanter Literatur zu Tage fördern kann und eine Überprüfung all dieser Literatur praktisch nicht umsetzbar ist. Zum anderen besteht die Gefahr, das relevante Literatur aufgrund eines zu allgemein gehaltenen Such-Strings nicht gefunden wird, so Napoleão et al. [649]. Unser erster Such-String lautete wie folgt:

> (*explainability* ∧ *privacy*) ∧ (*usability* ∨ *trustworthiness* ∨ „*privacy explanation*“ ∨ „*explainable privacy*“) ∧ (*-ai -xai -machine -neural -recommender*)

Um zu vermeiden, das Literatur aus den Bereichen XAI, ML oder Recommender-Systemen gefunden wird, schlossen wir derartige Begriffe aus. Insgesamt führte die Suche zu 84 Treffern, von denen wir fünf als relevant selektieren konnten.

Ziel der zweiten Suche war, Erkenntnisse über den Stand der Privatsphäreerklärungen und die Dokumente, die sie enthalten können, zu gewinnen. Spezielle Erklärungen zum Datenschutz sind i.d.R. nicht „allein stehend" zu finden, sondern meist in Dokumenten wie Datenschutzerklärungen oder Datenschutzhinweisen enthalten. Der Such-String umfasst daher auch diese Begriffe. Da das Ziel dieser Dokumente darin besteht, die informierte Zustimmung der Endnutzer einzuholen, wird auch die Zustimmung in die Suche einbezogen. Folglich wurde der zweite Such-String wie folgt definiert:

> **allintitle:** *privacy* ∧ (*explanation* ∨ *statement* ∨ *notice* ∨ *consent* ∨ *notification*) ∧ (*-ai -xai -machine -neural -recommender*)

Wir haben hier das Schlüsselwort allintitle genutzt, da die Suche sonst in über drei Millionen Treffern mündete. Mit dem Schlüsselwort erhielten wir 173 Treffer, von denen wir 13 als relevant markierten. Wir haben uns für eine dritte Suche entschieden, mit der wir gezielt den Sektor der mobilen Anwendungen abdecken wollten, da mobile Anwendungen heutzutage ein wesentlicher Bestandteil im Alltag der Gesellschaft darstellen [650]. Auch hier haben wir das Schlüsselwort allintitle im Suchstring verwendet, um eine hohe Relevanz potentieller Publikationskandidaten zu erhalten. Der Such-String wurde wie folgt definiert:

> **allintitle:** *app* ∧ *privacy* ∧ (*mobile* ∨ *smartphone* ∧ (*-covid*)

Da die LR zu Zeiten der Covid-19 Pandemie durchgeführt wurde und wir zu dieser Zeit sehr viele Publikationen im Bereich Smartphones und Privatsphäre erschienen sind, haben wir diese aus unserer Suche ausgeschlossen. Ein vorige Analyse der Ergebnisse ohne das Flag „-covid" offenbarte, das ein Großteil der Arbeiten für unsere LR keine Relevanz aufwiesen. Die dritte Datenbanksuche resultierte schließlich in 123 Treffern, von denen wir 24 für unsere LR auswählten. Folgende 42 Publikationen haben wir bei der Datenbanksuche gefunden: [651, 652, 653, 393, 399, 654, 655, 287, 656, 657, 658, 659, 279, 660, 661, 662, 663, 664, 665, 666, 667, 394, 668, 191, 396, 669, 405, 670, 671, 672, 673, 274, 290, 116, 674, 650, 268, 392, 49, 289, 675, 411]

E.1.4 Einschluss- und Ausschlusskriterien

IC_1 Adressiert mindestens eines der folgenden Themenbereiche:

- Eine mit Erklärbarkeit in beziehungstehende NFR wie z.B. Privacy, Trust, Trustworthiness, …
- Privacy Explanations (oder verwandte Konzepte)
- Dokumente (wie z.B. DSEs), die Erklärungen zur Privatsphäre enthalten

IC_3 *Peer-reviewed* Beitrag (Journal, Konferenz oder Workshop)

Wir haben Publikationen ausgeschlossen, die mindestens eins der folgenden Ausschlusskriterien erfüllten:

EC_1 nicht in englischer Sprache

EC_2 Tutorials, Proposals oder andere *non-peer-reviewed* Publikationen

EC_3 Arbeiten, die keinen Bezug zwischen Qualitätsaspekten sowie Privatsphäre und Erklärbarkeit herstellen

E.2 Prototyp der Konzeptstudie

E.2.1 Screenshots des Prototyps

Nachfolgend sind die verschiedenen Typen von Erklärungen der Konzeptstudie am Beispiel der präzisen Standortermittlung abgebildet.

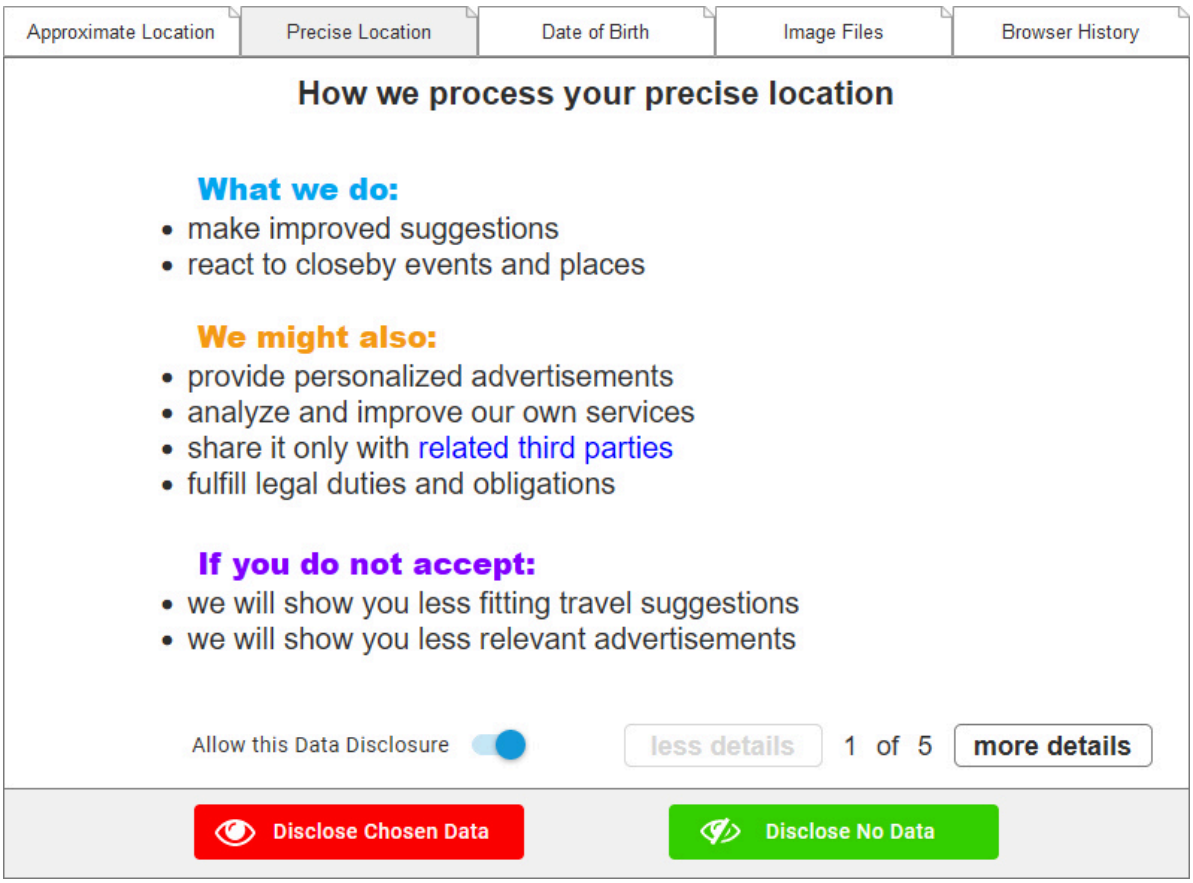


Abbildung E.1: Base-line Explanation für präzise Standortermittlung

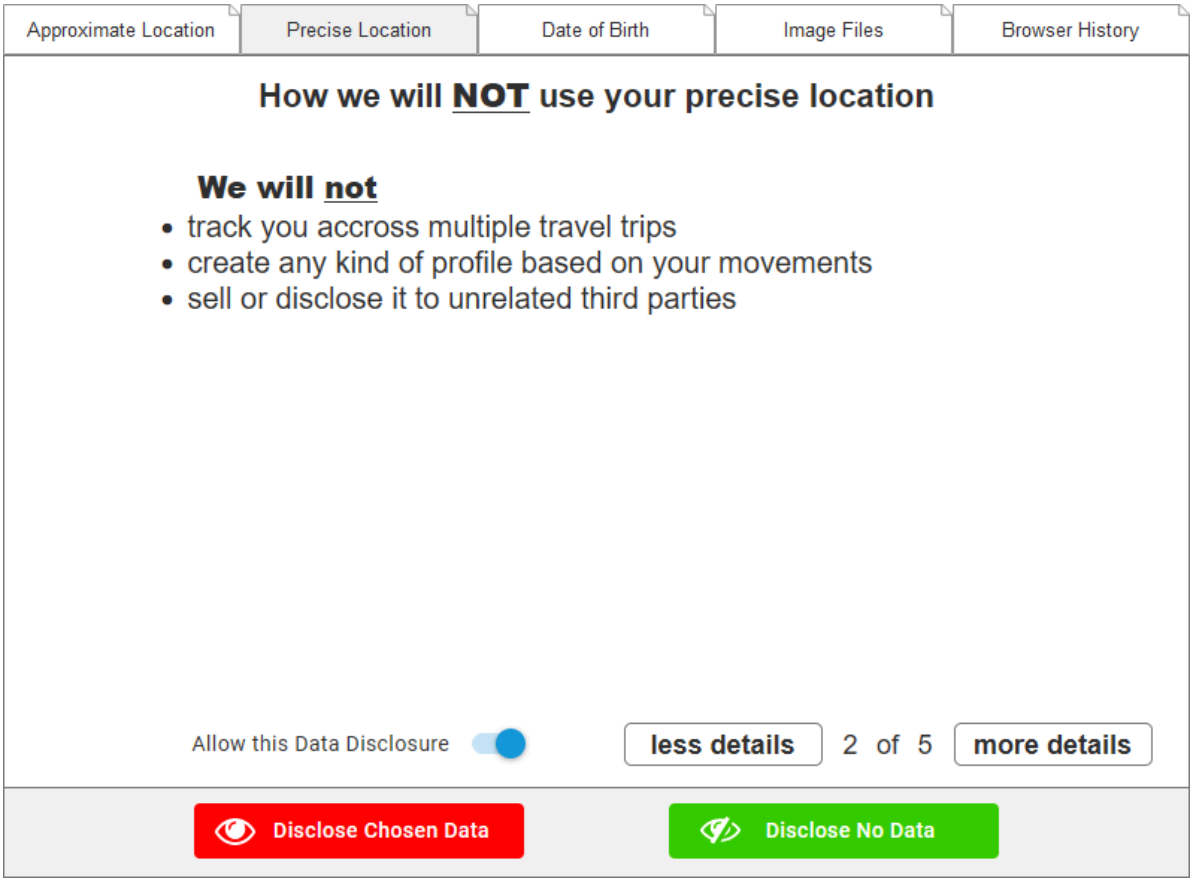


Abbildung E.2: Contrastive Explanation für präzise Standortermittlung

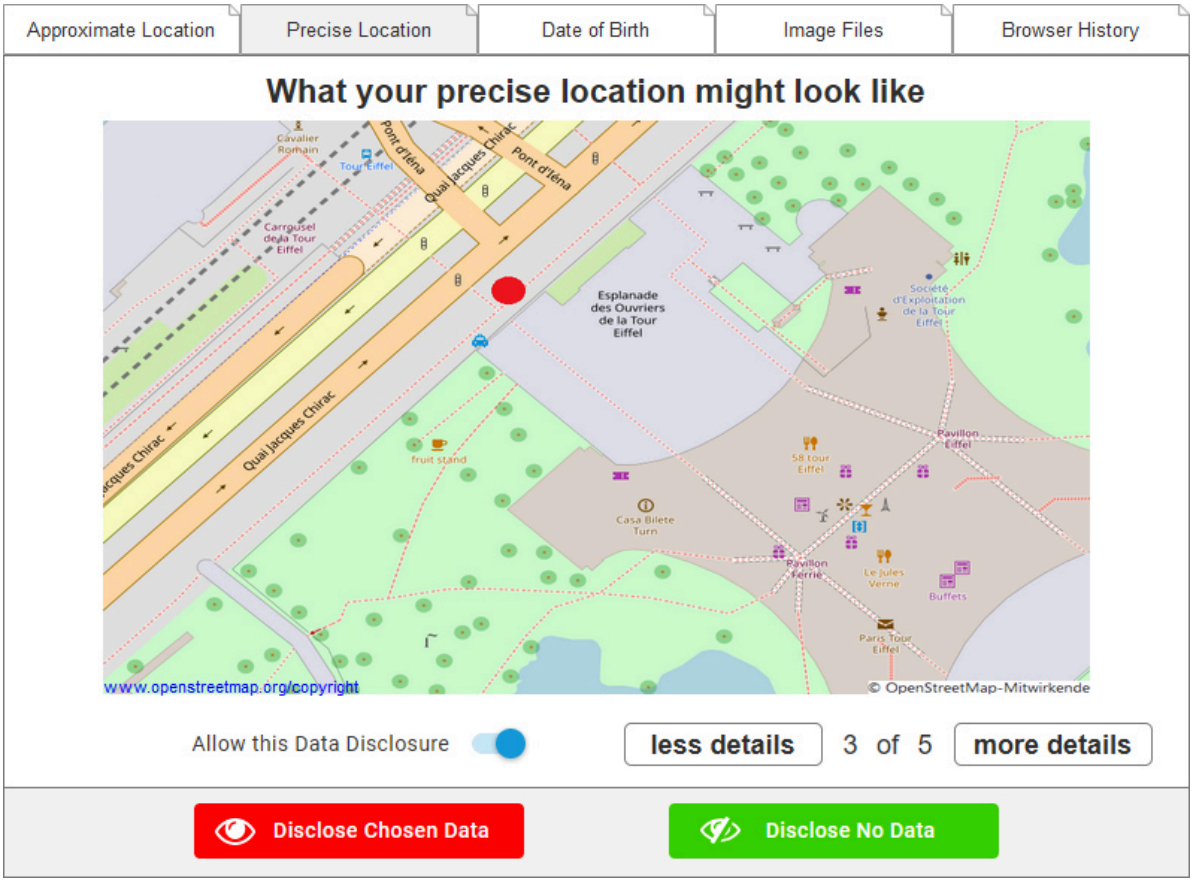


Abbildung E.3: Example-based Explanation für präzise Standortermittlung

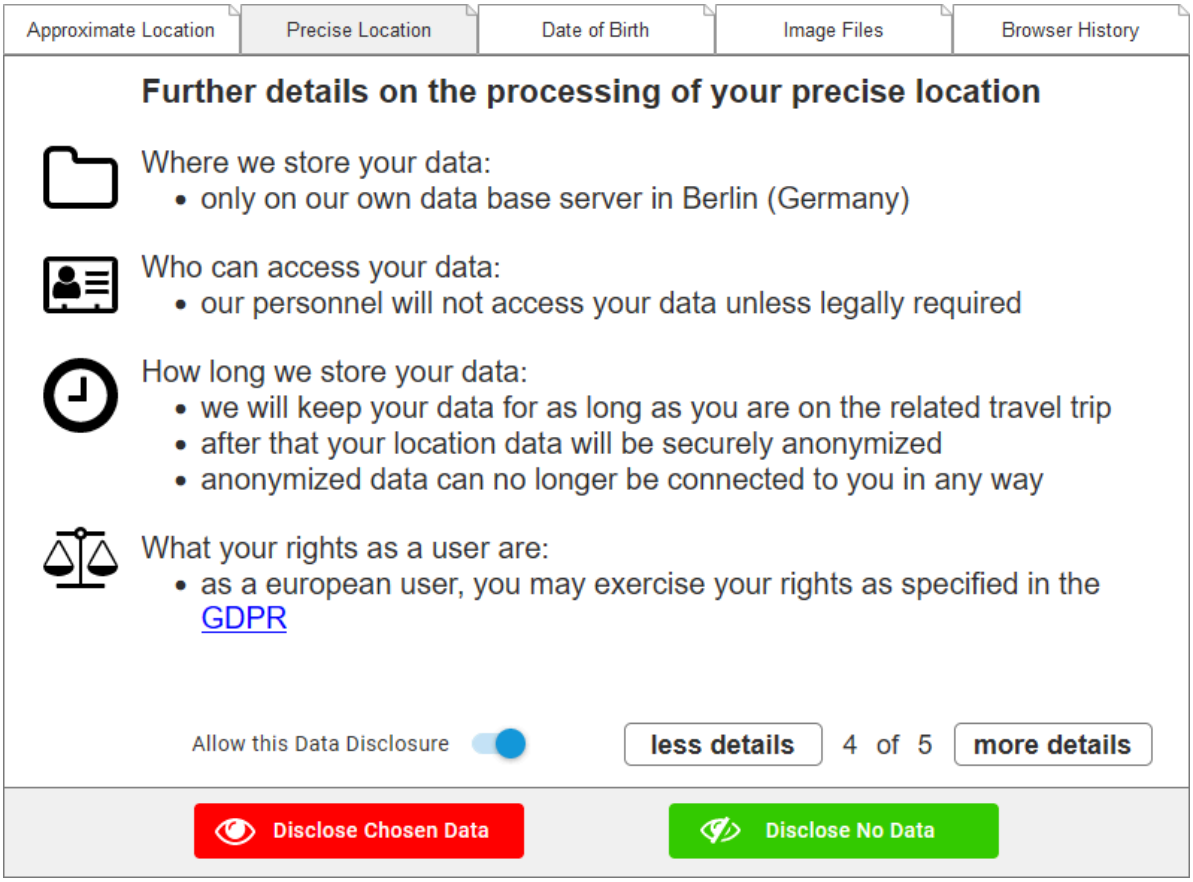


Abbildung E.4: Details Explanation für präzise Standortermittlung

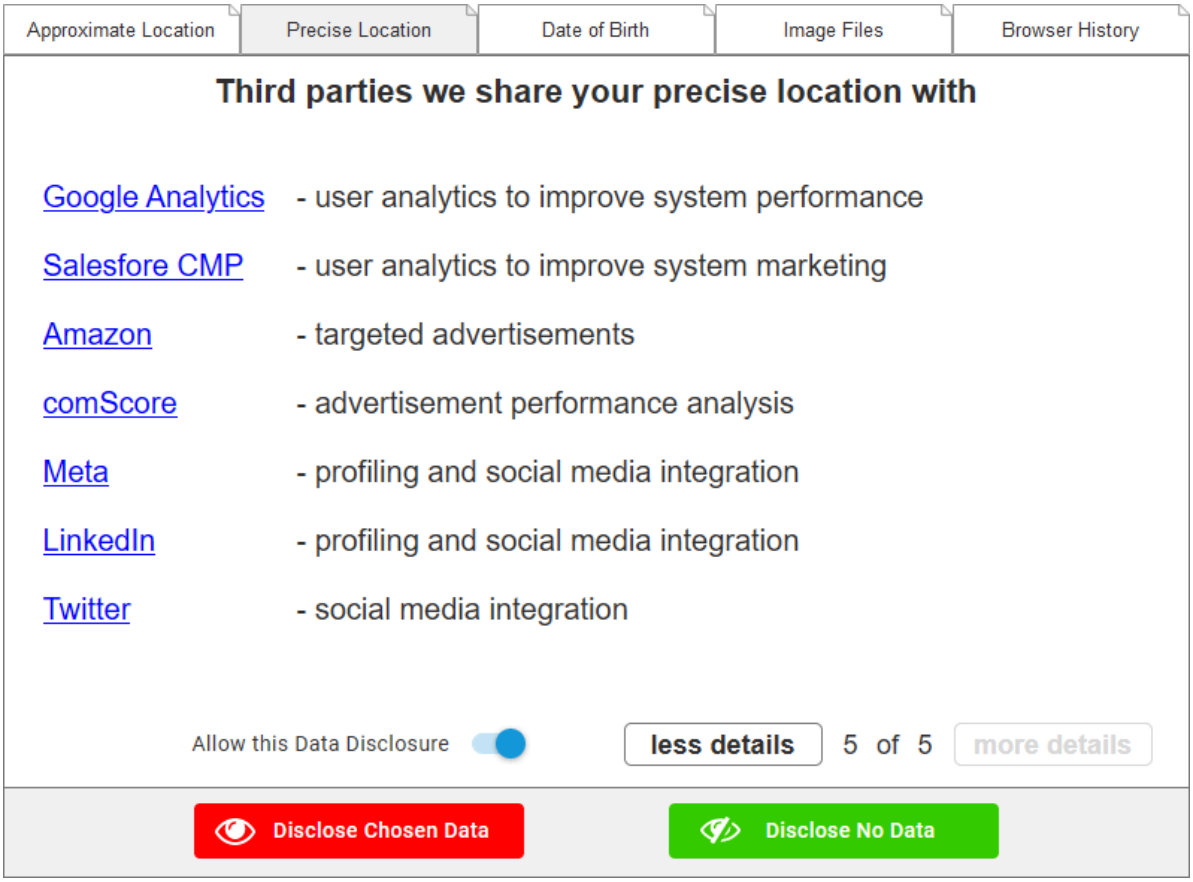
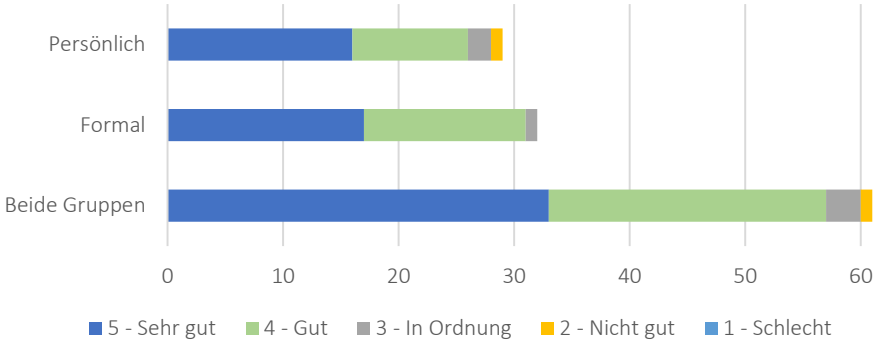


Abbildung E.5: Third Parties Explanation für präzise Standortermittlung

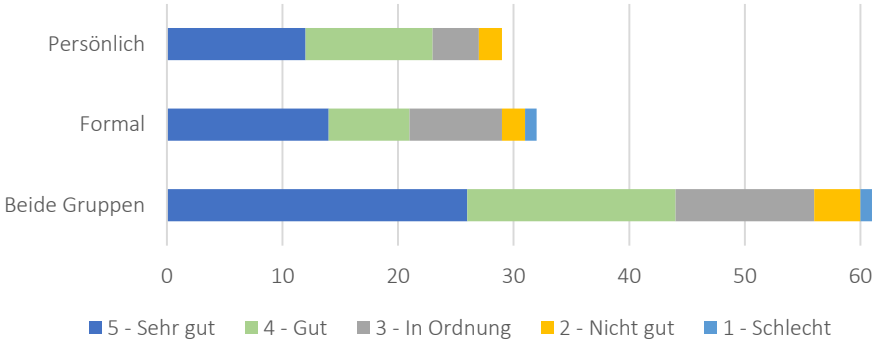
E.3 Evaluation der Konzeptstudie

Für den Umgang mit dem Prototyp wurden den Teilnehmern folgendes an [41] angelehntes Szenario an die Hand gegeben.

> „Stellen Sie sich folgendes Szenario vor: Sie wollen in ihrem Urlaub auf eine Reise ins Ausland gehen (z. B. nach Paris). Sie haben im Vorfeld recherchiert und haben dabei eine Reiseführer-App gefunden, die Sie verwenden wollen. Die App soll Ihnen Vorschläge für Sehenswürdigkeiten oder Ereignisse machen, die sich in Ihrer Nähe befinden. Als Sie die App aus dem Appstore installieren wollen, öffnet sich eine Privatsphäreerklärung (siehe Software-Prototyp). Mithilfe der Tabs können Sie zwischen den Daten-Arten wechseln. Mit ‚mehr Details' oder ‚weniger Details' wechseln Sie zwischen den Erklärungen. Navigieren Sie den Prototyp zunächst frei nach Interesse. Danach gibt es einige kleine Aufgaben darin zu erfüllen" [5].

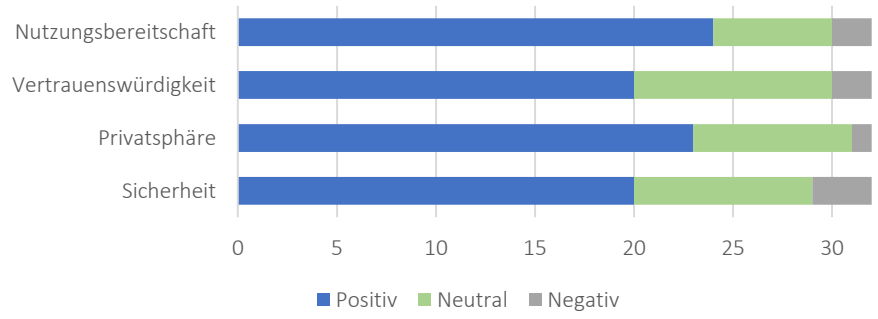


(a) Allgemeines Gefallen der Privacy Explanations

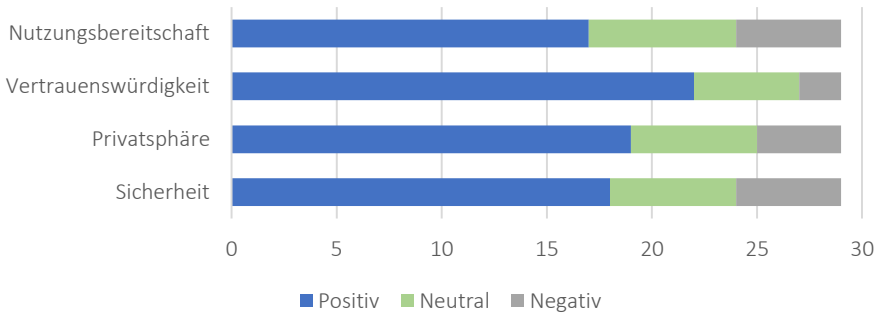


(b) Allgemein empfundene Glaubwürdigkeit der Privacy Explanations

Abbildung E.6: Gefallen und Glaubwürdigkeit der Privacy Explanations



(a) Überblick der Einflüsse von Privacy Explanations (Gruppe formal)



(b) Überblick der Einflüsse von Privacy Explanations (Gruppe Direkte Ansprache)

Abbildung E.7: Überblick Einflüsse von Privacy Explanations

E.3.1 Codes aus den Antworten der Teilnehmer

Tabelle E.1: Privacy Explanation Präferenzen der Teilnehmer

Code	Anzahl	Teilnehmer IDs
Tell ”What if not”	6	8, 35, 50, 55, 57, 59
Address users’ worries	18	2, 13, 16, 22, 26, 27, 29, 31, 35, 45, 48, 49, 50, 52, 54, 56, 59, 61
Provide helpful examples	26	2, 4, 9, 12, 14, 21, 25, 27, 28, 29, 30, 33, 34, 36, 39, 40, 44, 45, 47, 49, 50, 51, 54, 57, 60, 61
Do not necessarily provide contrast	11	10, 11, 15, 17, 20, 21, 34, 41, 43, 51, 55
Do not necessarily provide examples	9	1, 3, 8, 17, 19, 20, 31, 35, 46
Do not necessarily provide details	3	17, 19, 21
Provide PE as whole	15	14, 16, 19, 22, 27, 28, 30, 37, 46, 49, 50, 52, 53, 58, 59

E.4 Technischer Prototyp

E.4.1 Kontextszenarien

(i) Posten eines Bildes oder GIFs Das Posten von Bildmaterial ist ein fester Bestandteil bei der Interaktion mit anderen Nutzern in vielen sozialen Medien. Hierbei kann es sich um ein mit der eigenen Kamera aufgenommenes Bild handeln, was den Nutzer beispielsweise selbst abbildet oder ein Bild vom lokalen Speicher des Benutzers. Ein GIF stellt im allgemeinen Sinn natürlich ebenfalls ein Bild dar, jedoch nehmen diese animierten Bilder nochmals eine Sonderstellung bei Reaktion auf Posts etc. ein, ähnlich wie Emojis. Nicht nur das Hochladen eines Bildes, welches den Nutzer selbst oder andere zeigt, sondern auch das Posten eines Bildes/GIFs stellt einen Eingriff in die Privatsphäre dar, da es Stimmungen und Meinungen des Nutzers repräsentieren kann.

(ii) Angabe des eigenen Standorts Bei Twitter kann ein Nutzer im Zuge des Erstellens eines Posts anderen seinen Standort mitteilen. Technisch kann dies hardware-seitig über ein GPS-Modul erfolgen oder über die IP-Adresse des Nutzers. Twitter nutzt für die Geolokalisierung über die IP-Adresse den Drittanbieter Foursquare[1]. Zusätzlich nutzt Twitter die IP-Adresse seiner Nutzer, um diesen relevante Inhalte zu präsentieren. Beides stellt Eingriffe in die Privatsphäre dar, so dass der Nutzer informiert werden sollte.

(iii) Erstellung und Interaktion mit Posts Während der Interaktion mit Inhalten, erhebt und verarbeitet Twitter Nutzerdaten. Auskunft, um welche Daten es sich handelt, werden in der DSE von Twitter[2] erläutert. Über diese Datenpraktiken sollte ein Nutzer ebenfalls informiert werden, weshalb auch hier Privacy Explanations notwendig sind. Zur Interaktion zählen auch das Posten und Teilen von Inhalten. Hierbei werden Profilinformationen des Autors, wie beispielsweise Avatar und Benutzername etc. mit veröffentlicht, zusätzlich können weitere Daten mit Drittanbietern geteilt werden, je nach gewählten Privatsphäreeinstellungen des Nutzers.

(iv) Suchen von Inhalten gewählt Twitter bietet seine Nutzern eine Suchfunktion an. Laut der DSE werden die Suchanfragen gespeichert und ggf. weiterverarbeitet, so dass auch für dieses Kontextszenario eine Privacy Explanation notwendig ist.

[1] `https://location.foursquare.com/company/partners/`, zuletzt besucht am 05.03.2023
[2] `https://twitter.com/en/privacy`, zuletzt besucht am 05.03.2023

E.4.2 Aufgaben der Nutzerstudie

Entsprechend den vier Kontextszenarien gab es damit vier verbundene Aufgaben, die von den Studienteilnehmern bewältigt werden sollten: ① Posten eines Bildes und Mitteilung des eigenen Standorts, ② Kommentieren eines vorhandenen Posts mit einem GIF, ③ Einen existierenden Post mit einem Lesezeichen markieren und ④ Interaktion mit der Suche. Nur für die Aufgaben den Zeitmessungen erhielt die Experimentalgruppe den Prototyp mit Privacy Explanations und die Kontrollgruppe einen Prototyp ohne Privacy Explanations. Bei den übrigen Aufgaben verwendeten beide Gruppe den Prototyp mit Privacy Explanations.

Bei der Untersuchung hinsichtlich der Verständlichkeit der gezeigten Privacy Explanations wurde folgendes überprüft:

Tabelle E.2: Aussagen zur Überprüfung des Verständnisses der Privacy Explanations

Prüfung Privacy Explanation für Giphy
1. Die Erklärung enthält den Inhalt der Nutzungsbedingungen von Giphy
2. Die Erklärung beinhaltet, wie lange das gepostete GIF gespeichert wird
3. Die Erklärung beinhaltet, dass die eigenen Nutzerdaten an Giphy weitergeleitet werden
Prüfung Privacy Explanation für die Kamera
1. Die Erklärung enthält alternative Wege, um ein Bild zu posten
2. Die Erklärung beinhaltet den Inhalt der DSGVO
3. Die Erklärung enthält Informationen darüber, wann die Kamera angeschaltet wird

E.4.3 Bedienbarkeit des Systems mit Privacy Explanations

Um die Usability des Systems, bzw. genauer gesagt die Usability der präsentierten Privacy Explanations zu messen, haben wie Fragen des SUS genutzt. Es wurde sich für einen Auszug an Fragen entschieden, da es nur um einen Teilaspekt des Gesamtsystems ging und der Twitter-Klon als Ganzes bewertet werden sollte. Nachfolgend sind die gestellten Fragen aufgelistet:

Tabelle E.3: Gestellte Fragen aus dem SUS zur Messung der Usability der Privacy Explanations

Auszug SUS-Fragen
1. Ich kann mir vorstellen, Privacy Explanations regelmäßig zu nutzen
2. Ich denke, dass Privacy Explanations einfach zu nutzen sind
3. Ich denke, dass die Mehrheit an Personen Privacy Explanations nutzen können

E.4.4 Ergebnisse der Benutzerstudie - Relevanz der Privacy Explanations

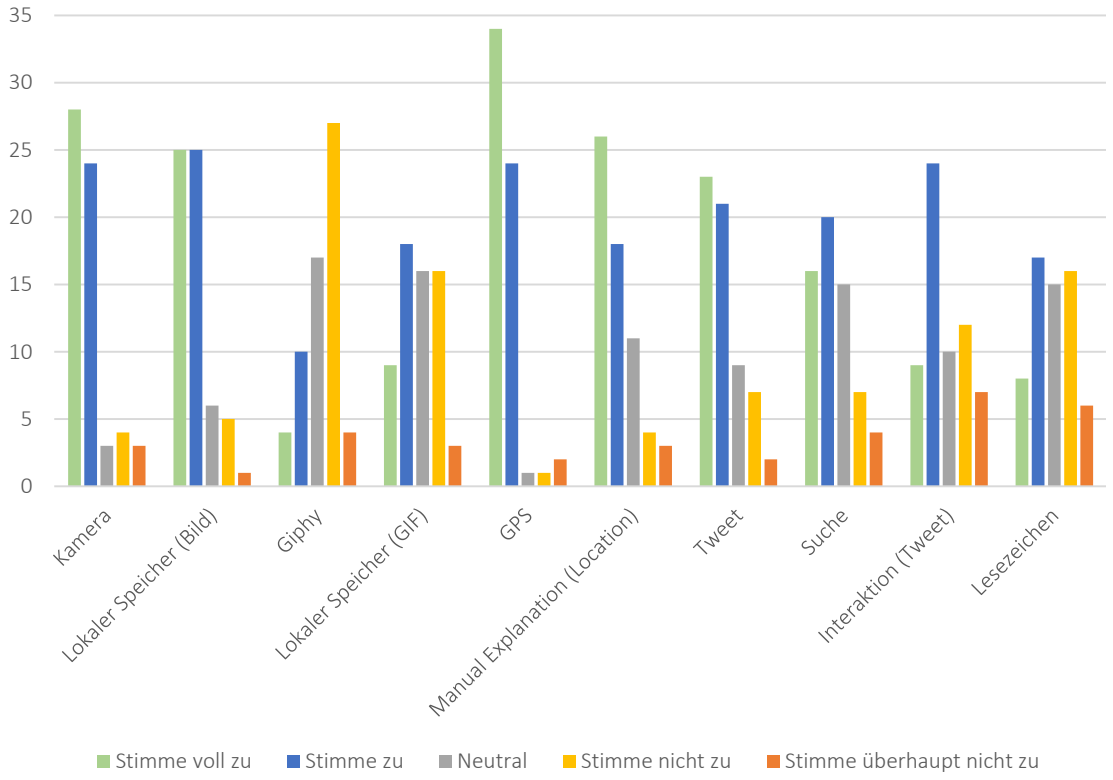
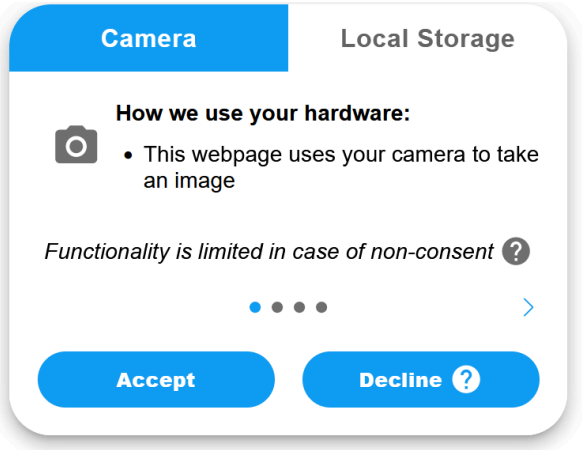


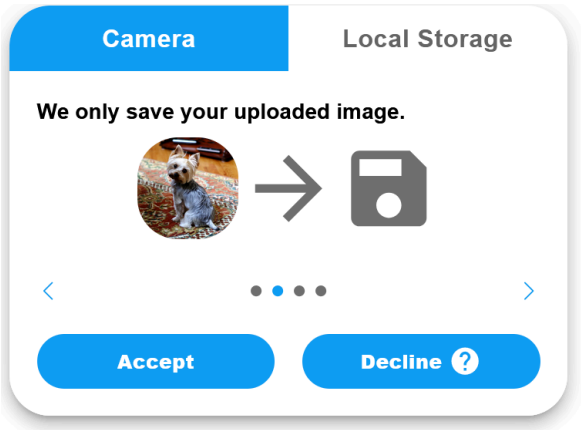
Abbildung E.8: Wahrgenommene Relevanz der Privacy Explanations

E.4.5 Screenshots

Nachfolgend sind exemplarisch Screenshots der Privacy Explanations für die Kameranutzung gezeigt.

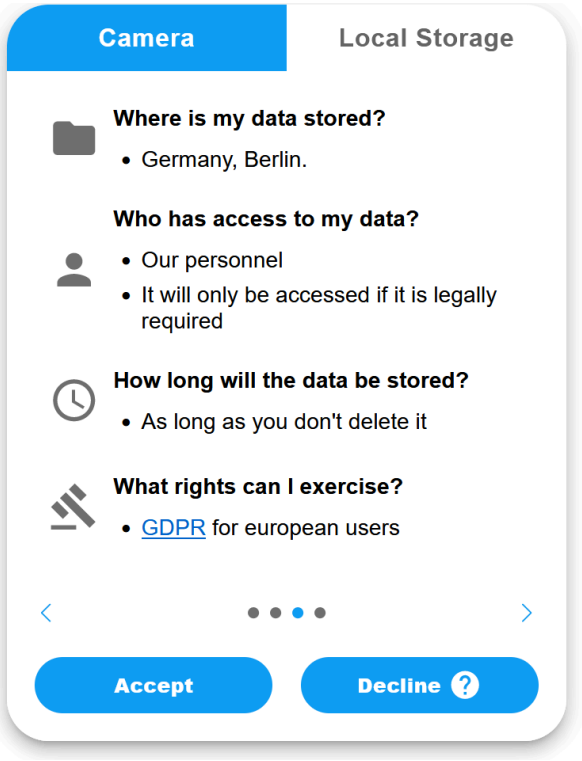


(a) Base-line Explanation für Kameranutzung

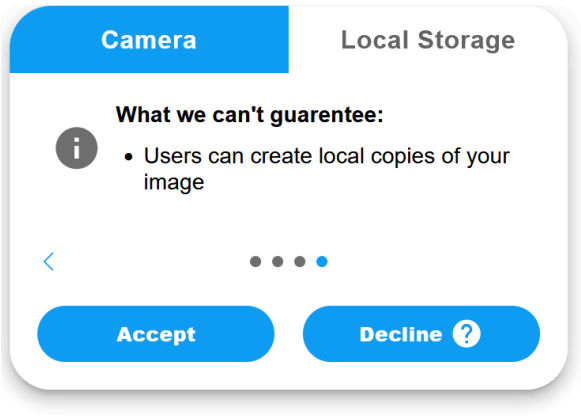


(b) Contrastive Explanation für Kameranutzung

Abbildung E.9: Screenshots 1/2



(a) Details Explanation für Kameranutzung



(b) Grenzen der Datenfreigabe für Kameranutzung

Abbildung E.10: Screenshots 2/2

Literaturverzeichnis

[1] E. Börger, B. Hörger, D. Parnas, and D. Rombach, "Requirements Capture, Documentation and Validation (Dagstuhl Seminar 99241)," Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, Dagstuhl Seminar Report 242, 1999. DOI: 10.4230/DagSemRep.242

[2] A. Hevner and S. March, "The information systems research cycle," *Computer*, vol. 36, no. 11, pp. 111–113, 2003. DOI: 10.1109/mc.2003.1244541

[3] L. Chazette, W. Brunotte, and T. Speith, "Explainable Software Systems: From Requirements Analysis to System Evaluation," *Requirements Engineering*, vol. 27, no. 4, pp. 457–487, Dec 2022. DOI: 10.1007/s00766-022-00393-5

[4] 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)*, 2021. DOI: 10.1109/RE51729.2021.00025 pp. 197–208.

[5] J. R. C. Droste, "Development of a Concept for Privacy Explanations and its Prototypical Evaluation," Master's thesis, Leibniz Universität Hannover, Fachgebiet Software Engineering, Hannover, Germany, April 2022.

[6] F. Volodarskis, "Konzeptionierung und prototypische Umsetzung von kontextuellen Privatsphäreerklärungen," Master's thesis, Leibniz Universität Hannover, Fachgebiet Software Engineering, Hannover, Germany, April 2023.

[7] J. A. Hölzing, *Die Kano-Theorie der Kundenzufriedenheitsmessung*. Wiesbaden, Deutschland: Gabler Verlag Wiesbaden, 2008.

[8] L. Köhler, "Automatisierte Analyse und visuelle Aufbereitung von Datenschutzerklärungen," Bachelor, Leibniz Universität Hannover, Fachgebiet Software Engineering, July 2021.

[9] A. Cavoukian *et al.*, "Privacy by Design: The 7 Foundational Principles," *Information and privacy commissioner of Ontario, Canada*, vol. 5, p. 12, 2010.

[10] F. G. of the United States, "Fair Information Practice Principles (FIPPs)," https://www.fpc.gov/resources/fipps/, 2023, Letzter Zugriff: 08.09.2023.

[11] A. Greenfield, *Everyware: The Dawning Age of Ubiquitous Computing*, ser. AIGA Design Press. Berkeley, CA, USA: Pearson Education, 2010.

[12] M. Janssen, M. Hartog, R. Matheus, A. Y. Ding, and G. Kuk, "Will Algorithms Blind People? The Effect of Explainable AI and Decision-Makers' Experience on AI-supported Decision-Making in Government," *Social Science Computer Review*, vol. 40, no. 2, pp. 478–493, 2022. DOI: 10.1177/0894439320980118

[13] J. Törnquist, "Computer-based decision support for railway traffic scheduling and dispatching: A review of models and algorithms," in *5th Workshop on Algorithmic Methods and Models for Optimization of Railways (ATMOS'05)*, ser. OpenAccess Series in Informatics (OASIcs), L. G. Kroon and R. H. Möhring, Eds., vol. 2. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2006. DOI: 10.4230/OASIcs.ATMOS.2005.659. ISSN 2190-6807

[14] S. Srinivasan and T. Kamalakannan, "Multi Criteria Decision Making in Financial Risk Management with a Multi-objective Genetic Algorithm," *Computational Economics*, vol. 52, no. 2, pp. 443–457, Aug 2018. DOI: 10.1007/s10614-017-9683-7

[15] J. Dräger and R. Müller-Eiselt, *Wir und die intelligenten Maschinen: Wie Algorithmen unser Leben bestimmen und wir sie für uns nutzen können*, 1st ed. München, Germany: Deutsche Verlags-Anstalt, 2019.

[16] Mike Walsh, "Welcome to the Algorithmic Age," 2018, Letzter Zugriff: 30.01.2023. [Online]. Available: https://www.mike-walsh.com/news/welcome-to-the-algorithmic-age

[17] B. Zheng, S. W. Yoon, and S. S. Lam, "Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms," *Expert Systems with Applications*, vol. 41, no. 4, Part 1, pp. 1476–1482, 2014. DOI: 10.1016/j.eswa.2013.08.044

[18] W. Sun, B. Zheng, and W. Qian, "Computer aided lung cancer diagnosis with deep learning algorithms," in *Medical Imaging 2016: Computer-Aided Diagnosis*, G. D. Tourassi and S. G. A. III, Eds., vol. 9785, International Society for Optics and Photonics. Spie, 2016. DOI: 10.1117/12.2216307 p. 97850z.

[19] M. K. Lee, D. Kusbit, E. Metsky, and L. Dabbish, "Working with Machines: The Impact of Algorithmic and Data-Driven Management on Human Workers," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ser. Chi '15. New York, NY, USA: Association for Computing Machinery, 2015. DOI: 10.1145/2702123.2702548 p. 1603–1612.

[20] Electronic Privacy Information Center, "AI in the Criminal Justice System," https://epic.org/issues/ai/ai-in-the-criminal-justice-system/, Letzter Zugriff: 19.01.2023.

[21] B. Stark, D. Stegmann, P. Jürgens, and M. Magin, "Are algorithms a threat to democracy? The rise of intermediaries: A challenge for public discourse," in *Governing Platforms*. Algorithmic Watch, 2020.

[22] M. N. Ndlela, "Social Media Algorithms, Bots and Elections in Africa," in *Social Media and Elections in Africa, Volume 1: Theoretical Perspectives and Election Campaigns*, M. N. Ndlela and W. Mano, Eds. Cham: Springer International Publishing, 2020, pp. 13–37. DOI: 10.1007/978-3-030-30553-6_2

[23] J. Hinds, E. J. Williams, and A. N. Joinson, ""It wouldn't happen to me": Privacy concerns and perspectives following the Cambridge Analytica scandal," *International Journal of Human-Computer Studies*, vol. 143, p. 102498, 2020. DOI: 10.1016/j.ijhcs.2020.102498

[24] M. Hu, "Cambridge Analytica's black box," *Big Data & Society*, vol. 7, no. 2, p. 7, 2020. DOI: 10.1177/2053951720938091

[25] D. Parkins, "The world's most valuable resource is no longer oil, but data," *The Economist*, vol. 6, 2017. [Online]. Available: https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data

[26] J. Wieringa, P. Kannan, X. Ma, T. Reutterer, H. Risselada, and B. Skiera, "Data analytics in a privacy-concerned world," *Journal of Business Research*, vol. 122, pp. 915–925, 2021. DOI: 10.1016/j.jbusres.2019.05.005

[27] O. Rana and J. Weinman, "Data as a Currency and Cloud-Based Data Lockers," *IEEE Cloud Computing*, vol. 2, no. 2, pp. 16–20, 2015. DOI: 10.1109/mcc.2015.46

[28] V. T. Patil and R. K. Shyamasundar, "Is Privacy a Myth for Facebook Users?" in *Proceedings of the 16th International Joint Conference on e-Business and Telecommunications, ICETE 2019 - Volume 2: SECRYPT, Prague, Czech Republic, July 26-28, 2019*, M. S. Obaidat and P. Samarati, Eds. SciTePress, 2019. DOI: 10.5220/0008018805100516 pp. 510–516.

[29] N. Wessels, A. Laubach, and P. Buxmann, "Personenbezogene Daten in der digitalen Ökonomie – Eine wirtschaftliche und juristische Betrachtung," in *Die Zukunft der Datenökonomie: Zwischen Geschäftsmodell, Kollektivgut und Verbraucherschutz*, C. Ochs, M. Friedewald, T. Hess, and J. Lamla, Eds. Wiesbaden: Springer Fachmedien Wiesbaden, 2019, pp. 11–27. DOI: 10.1007/978-3-658-27511-2_2

[30] T. Dinev, "Why would we care about privacy?" *European Journal of Information Systems*, vol. 23, no. 2, pp. 97–102, Mar 2014. DOI: 10.1057/ejis.2014.1

[31] B. Schneier, *Data and Goliath: The hidden battles to collect your data and control your world*, 1st ed. New York, NY, USA: W. W. Norton & Company, 2015.

[32] S. Garcia-Rivadulla, "Personalization vs. Privacy: An Inevitable Trade-off?" *IFLA Journal*, vol. 42, no. 3, pp. 227–238, 2016. DOI: 10.1177/0340035216662890

[33] C. Tun-Min, J. King, and N. J. King, "Privacy versus reward: Do loyalty programs increase consumers' willingness to share personal information with third-party advertisers and data brokers?" *Journal of Retailing and Consumer Services*, vol. 28, pp. 296–303, 2016. DOI: 10.1016/j.jretconser.2015.01.005

[34] T. Barnett White, "Consumer Disclosure and Disclosure Avoidance: A Motivational Framework," *Journal of Consumer Psychology*, vol. 14, no. 1, pp. 41–51, 2004.

[35] J. B. Earp and D. Baumer, "Innovative Web Use to Learn about Consumer Behavior and Online Privacy," *Commun. ACM*, vol. 46, no. 4, p. 81–83, apr 2003. DOI: 10.1145/641205.641209

[36] I.-H. Hann, K.-L. Hui, T. Lee, and I. Png, "Online Information Privacy: Measuring the Cost-Benefit Trade-Off," in *International Conference on Information Systems (ICIS)*, 2002, pp. 1–11.

[37] G. Iachello and J. Hong, *End-User Privacy in Human-Computer Interaction, Foundation and Trends® in Human– Computer Interaction*, 1st ed., B. Bederson, G. Abowd, J. Grudin, C. Lewis, J. Nielsen, D. Norman, D. Olsen, G. Olson, and S. Oviatt, Eds. Hanover, MA, USA: now Publishers Inc., 2007.

[38] C. Utz, M. Degeling, S. Fahl, F. Schaub, and T. Holz, "(Un)Informed Consent: Studying GDPR Consent Notices in the Field," in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, ser. Ccs '19. New York, NY, USA: Association for Computing Machinery, 2019. DOI: 10.1145/3319535.3354212 p. 973–990.

[39] Bundesamt für Sicherheit in der Informationstechnik, "Cookie - Glossareintrag," https://www.bsi.bund.de/SharedDocs/Glossareintraege/DE/C/Cookie.html, Letzter Zugriff: 20.01.2023.

[40] T. H. Soe, O. E. Nordberg, F. Guribye, and M. Slavkovik, "Circumvention by Design - Dark Patterns in Cookie Consent for Online News Outlets," in *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–12.

[41] W. Brunotte, A. Specht, L. Chazette, and K. Schneider, "Privacy Explanations – A Means to End-User Trust," *Journal of Systems and Software*, vol. 195, p. 111545, 2023. DOI: 10.1016/j.jss.2022.111545

[42] W. Brunotte, L. Chazette, L. Köhler, J. Klunder, and K. Schneider, "What About My Privacy? Helping Users Understand Online Privacy Policies," in *Proceedings of the International Conference on Software and System Processes and International Conference on Global Software Engineering*, ser. Icssp'22. New York, NY, USA: Association for Computing Machinery, 2022. DOI: 10.1145/3529320.3529327 p. 56–65.

[43] F. H. Cate, "The Limits of Notice and Choice," *IEEE Security Privacy*, vol. 8, no. 2, pp. 59–62, 2010.

[44] President's Concil of Advisors on Science and Technology, "Big data and privacy: A technological perspective," Executive Office of the President, Tech. Rep., May 2014, report to the president.

[45] C. Jensen and C. Potts, "Privacy Policies as Decision-Making Tools: An Evaluation of Online Privacy Notices," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. Chi '04. New York, NY, USA: Association for Computing Machinery, 2004. DOI: 10.1145/985692.985752 p. 471–478.

[46] A. M. McDonald and L. F. Cranor, "The Cost of Reading Privacy Policies 2008 Privacy Year in Review," *I/S: A Journal of Law and Policy for the Information Society*, vol. 4, no. 3, pp. 543–568, 2009 2008.

[47] I. Pollach, "What's Wrong with Online Privacy Policies?" *Commun. ACM*, vol. 50, no. 9, p. 103–108, 2007.

[48] J. R. Reidenberg, T. Breaux, L. F. Cranor, B. French, A. Grannis, J. T. Graves, F. Liu, A. McDonald, T. B. Norton, R. Ramanath, N. C. Russell, N. Sadeh, and F. Schaub, "Disagreeable Privacy Policies: Mismatches Between Meaning and Users' Understanding," *Berkeley Technology Law Journal*, vol. 30, no. 1, pp. 39–88, 2015.

[49] A. E. Waldman, "Privacy, Notice, and Design ," *Stanford Technology Law Review*, vol. 21, no. 1, pp. 74–127, 2018.

[50] R. Cummings, G. Kaptchuk, and E. M. Redmiles, "Ï 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*, ser. Ccs '21. New York, NY, USA: Association for Computing Machinery, 2021. DOI: 10.1145/3460120.3485252 p. 3037–3052.

[51] J. Phelps, G. Nowak, and E. Ferrell, "Privacy Concerns and Consumer Willingness to Provide Personal Information," *Journal of Public Policy & Marketing*, vol. 19, no. 1, pp. 27–41, 2000. DOI: 10.1509/jppm.19.1.27.16941

[52] M. A. Köhl, K. Baum, M. Langer, D. Oster, T. Speith, and D. Bohlender, "Explainability as a Non-Functional Requirement," in *27th IEEE International Requirements Engineering Conference (RE)*. Ieee, 2019. DOI: 10.1109/re.2019.00046 pp. 363–368.

[53] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, *Experimentation in Software Engineering*. Heidelberg, Deutschland: Springer Science & Business Media, 2012.

[54] V. Basili and H. Rombach, "The TAME project: towards improvement-oriented software environments," *IEEE Transactions on Software Engineering*, vol. 14, no. 6, pp. 758–773, 1988. DOI: 10.1109/32.6156

[55] J. Nielsen, *Usability Engineering*, 1st ed. San Francisco, CA, USA: Morgan Kaufmann, November 1994.

[56] L. Chazette, "Requirements Engineering for Explainable Systems," Ph.D. dissertation, Leibniz University Hannover, 2022. DOI: 10.15488/13261

[57] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design Science in Information Systems Research," *MIS Quarterly*, vol. 28, no. 1, pp. 75–105, 2004. DOI: 10.2307/25148625

[58] J. S. Valacich and C. Schneider, *Information Systems Today: Managing the Digital World, Global Edition*, 8th ed. New York, New Yor, USA: Pearson, 2018.

[59] I. Sommerville, *Software Engineering*, 8th ed. Addison-Wesley, 2007.

[60] The Institute of Electrical and Electronics Engineers, "IEEE Recommended Practice for Architectural Description for Software-Intensive Systems," *IEEE Std 1471-2000*, pp. 1–30, Oct 2000. DOI: 10.1109/IEEESTD.2000.91944

[61] R. E. Freeman, *Strategic management: A stakeholder approach*. Cambridge university press, 2010.

[62] G. Kotonya and I. Sommerville, *Requirements Engineering: Processes and Techniques*. Wiley Publishing, 1998.

[63] A. Pouloudi, "Aspects of the Stakeholder Concept and their implications for Information Systems Development," in *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences. 1999. HICSS-32. Abstracts and CD-ROM of Full Papers*. IEEE, 1999, pp. 17–pp.

[64] H. Sharp, A. Finkelstein, and G. Galal, "Stakeholder Identification in the Requirements Engineering Process," in *Proceedings. Tenth International Workshop on Database and Expert Systems Applications. DEXA 99*. Ieee, 1999, pp. 387–391.

[65] J. McManus, *Managing Stakeholders in Software Development Projects*. Butterworth-Heinemann, 2007.

[66] L. C. Ballejos and J. M. Montagna, "Method for Stakeholder Identification in Inter-organizational Environments," *Requirements engineering*, vol. 13, no. 4, pp. 281–297, 2008.

[67] M. Glinz and R. J. Wieringa, "Guest Editors' Introduction: Stakeholders in Requirements Engineering," *IEEE Software*, vol. 24, no. 2, pp. 18–20, 2007. DOI: 10.1109/ms.2007.42

[68] P. Zave, "Classification of Research Efforts in Requirements Engineering," *ACM Comput. Surv.*, vol. 29, no. 4, p. 315–321, dec 1997. DOI: 10.1145/267580.267581

[69] B. Nuseibeh and S. Easterbrook, "Requirements Engineering: A Roadmap," in *Proceedings of the Conference on The Future of Software Engineering*, ser. ICSE '00. New York, NY, USA: Association for Computing Machinery, 2000. DOI: 10.1145/336512.336523 p. 35–46.

[70] M. Glinz, "A glossary of requirements engineering terminology," *Standard Glossary of the Certified Professional for Requirements Engineering (CPRE) Studies and Exam, Version 1.7*, vol. 1.7, p. 130, 2017.

[71] K. Pohl and C. Rupp, *Requirements Engineering; Fundamentals, Principles, and Techniques*, 1st ed. Heidelberg, Germany: Springer Berlin, Heidelberg, 2010.

[72] K. Pohl and C. Rupp, *Basiswissen Requirements Engineering: Aus- und Weiterbildung nach IREB-Standard zum Certified Professional for Requirements Engineering Foundation Level*, 5th ed. Heidelberg, Germany: dpunkt.verlag, 2021.

[73] C. Rupp and die SOPHISTen, *Requirements-Engineering und-Management: Aus der Praxis von klassisch bis agil*, 6th ed. Carl Hanser Verlag München, 2014.

[74] J. Doerr, T. Koenig, T. Olsson, and S. Adam, "Das ReqMan Prozessrahmenwerk," *IESE-Report Nr. 141.06/D*, 2006.

[75] J. J. Carr, "Requirements engineering and management: the key to designing quality complex systems," *The TQM Magazine*, vol. 12, no. 6, pp. 400–407, Jan 2000. DOI: 10.1108/09544780010351760

[76] T. Hall, S. Beecham, and R. A., "Requirements problems in twelve software companies: an empirical analysis," *IEE Proceedings - Software*, vol. 149, pp. 153–160(7), October 2002.

[77] G. N. Aranda, A. Vizcaíno, and M. Piattini, "A framework to improve communication during the requirements elicitation process in GSD projects," *Requirements Engineering*, vol. 15, no. 4, pp. 397–417, Nov 2010. DOI: 10.1007/s00766-010-0105-9

[78] C. Werner, Z. S. Li, N. Ernst, and D. Damian, "The Lack of Shared Understanding of Non-Functional Requirements in Continuous Software Engineering: Accidental or Essential?" in *2020 IEEE 28th International Requirements Engineering Conference (RE)*, 2020. DOI: 10.1109/RE48521.2020.00021 pp. 90–101.

[79] M. Glinz and S. A. Fricker, "On shared understanding in software engineering: an essay," *Computer Science - Research and Development*, vol. 30, no. 3, pp. 363–376, Aug 2015. DOI: 10.1007/s00450-014-0256-x

[80] N. Donald A., *The Design Of Everday Things*. New York, NY, USA: Basic Books, 2013.

[81] S. Easterbrook, "Coordination breakdowns: why groupware is so difficult to design," in *Proceedings of the Twenty-Eighth Annual Hawaii International Conference on System Sciences*, vol. 4, 1995. DOI: 10.1109/HICSS.1995.375730 pp. 191–199 vol.4.

[82] K. Schneider, *Abenteuer Softwarequalität: Grundlagen und Verfahren für Qualitätssicherung und Qualitätsmanagement*. Heidelberg, DE: dpunkt.verlag, 2012.

[83] P. T. Devanbu and S. Stubblebine, "Software Engineering for Security: A Roadmap," in *Proceedings of the Conference on The Future of Software Engineering*, ser. ICSE '00. New York, NY, USA: Association for Computing Machinery, 2000. DOI: 10.1145/336512.336559 p. 227–239.

[84] H. Krasner, "The Cost of Poor Software Quality in the US: A 2020 Report," *CISQ Consortium for Information & Software Quality*, 2021.

[85] A. Gillies, *Software Quality: Theory and Management*, 3rd ed. Lulu.com, 2011.

[86] P. J. Denning, "Editorial: What is Software Quality?" *Commun. ACM*, vol. 35, no. 1, p. 13–15, jan 1992. DOI: 10.1145/129617.384272

[87] D. A. Garvin, "What Does „Product Quality" Really Mean?" *Sloan Management Review*, vol. 25, pp. 25–43, 1984.

[88] B. Kitchenham and S. Pfleeger, "Software Quality: The Elusive Target [special issues section]," *IEEE Software*, vol. 13, no. 1, pp. 12–21, 1996. DOI: 10.1109/52.476281

[89] The Institute of Electrical and Electronics Engineers, "IEEE Standard for Software Quality Assurance Processes," *IEEE Std 730-2014 (Revision of IEEE Std 730-2002)*, pp. 1–138, 2014. DOI: 10.1109/IEEESTD.2014.6835311

[90] J. Mylopoulos, L. Chung, and B. Nixon, "Representing and using nonfunctional requirements: a process-oriented approach," *IEEE Transactions on Software Engineering*, vol. 18, no. 6, pp. 483–497, 1992. DOI: 10.1109/32.142871

[91] M. Glinz, "On Non-Functional Requirements," in *15th IEEE International Requirements Engineering Conference (RE 2007)*, 2007. DOI: 10.1109/RE.2007.45 pp. 21–26.

[92] K. Bessiere, I. Ceaparu, J. Lazar, J. Robinson, and B. Shneiderman, "Understanding Computer User Frustration: Measuring and Modeling the Disruption from Poor Designs," University of Maryland & Towson University, Maryland, USA, Tech. Rep., 2002, https://hdl.handle.net/1903/1233.

[93] B. C. V. Fraassen, "The Pragmatics of Explanation," *American Philosophical Quarterly*, vol. 14, no. 2, pp. 143–150, 1977.

[94] D. Sandborg, "Mathematical Explanation and the Theory of Why-Questions," *The British Journal for the Philosophy of Science*, vol. 49, no. 4, pp. 603–624, 1998.

[95] C. G. Hempel and P. Oppenheim, "Studies in the Logic of Explanation," *Philosophy of Science*, vol. 15, no. 2, pp. 135–175, 1948.

[96] R. C. Schank, "Explanation: A First Pass," in *Experience, Memory, and Reasoning*, 1st ed., J. L. Kolodner and C. K. Riesbeck, Eds. Hillsdale, Newy Jersey, USA: Lawrence Erlbaum Associates, 1986, ch. 9, pp. 139–165.

[97] C. Geminn and A. Roßnagel, "„Privatheit" und „Privatsphäre"aus der Perspektive des Rechts – ein Überblick," *JuristenZeitung*, vol. 70, no. 14, pp. 703–708, 2015.

[98] P. Grimm and H. Krah, "Privatsphäre," in *Handbuch Medien- und Informationsethik*, J. Heesen, Ed. Stuttgart: J.B. Metzler, 2016, pp. 178–185. DOI: 10.1007/978-3-476-05394-7_24

[99] H. Behrendt, W. Loh, T. Matzner, and C. Misselhorn, *Privatsphäre 4.0 - Eine Neuverortung des Privaten im Zeitalter der Digitalisierung*, 1st ed. Berlin, Germany: J.B. Metzler Stuttgart, 2019.

[100] K. Nissim and A. Wood, "Is privacy *privacy*?" *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 376, no. 2128, p. 20170358, 2018. DOI: 10.1098/rsta.2017.0358

[101] Y. N. Harari, *Sapiens: A Brief History of Humankind*, 1st ed. Dublin, Ireland: Vintage, Penguin Random House UK, April 2015.

[102] B. Rösler, *Der Wert des Privaten.* Frankfurt a.M., Germany: Suhrkamp, 2001.

[103] B. Rössler, "Anonymität und Privatheit," in *Anonymität im Internet: Grundlagen, Methoden und Tools zur Realisierung eines Grundrechts*, H. Bäumler and A. von Mutius, Eds. Wiesbaden: Vieweg+Teubner Verlag, 2003, pp. 27–40. DOI: 10.1007/978-3-663-05790-1_3

[104] A. F. Westin, *Privacy and Freedom.* ig Publishing, 2015.

[105] J. Rubenfeld, "The Right of Privacy," *Harvard Law Review*, vol. 102, no. 4, pp. 737–807, 1989.

[106] L. C. Velecky, "The Concept of Privacy," in *Privacy.* New York, NY, USA: John Wiley & Sons, 1978, ch. 2, pp. 13–34.

[107] S. M. Jourard, "Some Psychological Aspects of Privacy," *Law and Contemporary Problems*, vol. 31, no. 2, pp. 307–318, 1966.

[108] T. Dienlin and S. Trepte, "Is the Privacy Paradox a Relic of the Past? An in-depth Analysis of Privacy Attitudes and Privacy Behaviors," *European Journal of Social Psychology*, vol. 45, no. 3, pp. 285–297, 2015. DOI: 10.1002/ejsp.2049

[109] M. Rudolph, D. Feth, and S. Polst, "Why Users Ignore Privacy Policies – A Survey and Intention Model for Explaining User Privacy Behavior," in *Human-Computer Interaction. Theories, Methods, and Human Issues*, M. Kurosu, Ed. Cham: Springer International Publishing, 2018. DOI: 10.1007/978-3-319-91238-7_45 pp. 587–598.

[110] S. B. Barnes, "A privacy paradox: Social networking in the United States," *First Monday*, vol. 11, no. 9, Sep. 2006. DOI: 10.5210/fm.v11i9.1394

[111] N. Gerber, P. Gerber, and M. Volkamer, "Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior," *Computers & Security*, vol. 77, pp. 226–261, 2018. DOI: 10.1016/j.cose.2018.04.002

[112] S. Kokolakis, "Privacy Attitudes and Privacy Behaviour: A Review of Current Research on the Privacy Paradox Phenomenon," *Computers & Security*, vol. 64, pp. 122–134, 2017. DOI: 10.1016/j.cose.2015.07.002

[113] E. Hargittai and A. Marwick, "„What Can I Really Do?" Explaining the Privacy Paradox with Online Apathy," *International Journal of Communication*, vol. 10, no. 0, 2016.

[114] S. Pötzsch, "Privacy Awareness: A Means to Solve the Privacy Paradox?" in *The Future of Identity in the Information Society*, V. Matyáš, S. Fischer-Hübner, D. Cvrček, and P. Švenda, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 226–236.

[115] R. Bandara, M. Fernando, and S. Akter, "Explicating the privacy paradox: A qualitative inquiry of online shopping consumers," *Journal of Retailing and Consumer Services*, vol. 52, p. 101947, 2020. DOI: 10.1016/j.jretconser.2019.101947

[116] I. Pentina, L. Zhang, H. Bata, and Y. Chen, "Exploring privacy paradox in information-sensitive mobile app adoption: A cross-cultural comparison," *Computers in Human Behavior*, vol. 65, pp. 409–419, 2016.

[117] A. I. Anton, J. B. Earp, and J. D. Young, "How internet users' privacy concerns have evolved since 2002," *IEEE Security & Privacy*, vol. 8, no. 1, pp. 21–27, 2010. DOI: 10.1109/MSP.2010.38

[118] W. Brunotte, L. Chazette, and K. Korte, "Can Explanations Support Privacy Awareness? A Research Roadmap," in *2021 IEEE 29th International Requirements Engineering Conference Workshops (REW)*, 2021. DOI: 10.1109/rew53955.2021.00032 pp. 176–180.

[119] D. Zweig and J. Webster, "Where is the line between benign and invasive? An examination of psychological barriers to the acceptance of awareness monitoring systems," *Journal of Organizational Behavior*, vol. 23, no. 5, pp. 605–633, 2002. DOI: 10.1002/job.157

[120] E. McCallister, T. Grance, and K. A. Scarfone, "Guide to protecting the confidentiality of Personally Identifiable Information (PII)," National Institute of Standards and Technology, Gaithersburg, MD, Tech. Rep., 2010. DOI: 10.6028/NIST.SP.800-122

[121] F. Thouvenin and N. Braun Binder, "Transparenz durch Datenschutzerklärungen von Behörden," *Zeitschrift für Schweizerisches Recht (ZSR)*, pp. 5–29, 2022. DOI: 10.5167/uzh-231859

[122] European Parliament and Council, "General Data Protection Regulation," https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02016R0679-20160504, 2016, Letzter Zugriff: 16.03.2022.

[123] European Union, *Charter of Fundamental Rights of the European Union.* Brussels: European Union, 2010, vol. 53.

[124] G. G. Fuster and R. Gellert, "The Fundamental Right of Data Protection in the European Union: In Search of an Uncharted Right," *International Review of Law, Computers & Technology*, vol. 26, no. 1, pp. 73–82, 2012. DOI: 10.1080/13600869.2012.646798

[125] G. Miglicco, "GDPR is here and it is time to get serious," *Computer Fraud & Security*, vol. 2018, no. 9, pp. 9–12, 2018. DOI: 10.1016/S1361-3723(18)30085-X

[126] P. M. Schwartz, "Privacy and Democracy in Cyberspace ," *Vanderbilt Law Review*, vol. 52, p. 1607, 1999.

[127] United States Department of Justice, "The Privacy Act of 1974 – 2020 Edition," https://www.justice.gov/opcl/overview-privacy-act-1974-2020-edition, 2022, Letzter Zugriff: 30.12.2022.

[128] J. Beverage, "The Privacy Act of 1974: An Overview," *Duke Law Journal*, vol. 1976, no. 2, pp. 301–329, 1976.

[129] L. De la Torre, "A Guide to the California Consumer Privacy Act of 2018," in *Santa Clara University Legal Studies Research Paper*, 2018. DOI: 10.2139/ssrn.3275571

[130] E. Goldman, "An Introduction to the California Consumer Privacy Act (CCPA)," in *Santa Clara University Legal Studies Research Paper*, 2020. DOI: 10.2139/ssrn.3211013

[131] G. Stamenkov, "Genealogy of the Fair Information Practice Principles," *International Journal of Law and Management*, vol. 65, no. 3, pp. 242–260, Jan 2023. DOI: 10.1108/IJLMA-07-2022-0149

[132] ISO Central Secretary, "ISO/IEC 29100:2011 Information Technology — Security Techniques — Privacy Framework," International Organization for Standardization, Geneva, CH, Standard ISO/IEC 29100:2011(E), 2011. [Online]. Available: https://www.iso.org/standard/45123.html

[133] U. Sury, "Die unterbewertete Datenschutzerklärung," *Informatik Spektrum*, vol. 44, no. 6, pp. 459–460, Dec 2021. DOI: 10.1007/s00287-021-01416-1

[134] K.-W. Wu, S. Y. Huang, D. C. Yen, and I. Popova, "The Effect of Online Privacy Policy on Consumer Privacy Concern and Trust," *Computers in Human Behavior*, vol. 28, no. 3, pp. 889–897, 2012. DOI: https://doi.org/10.1016/j.chb.2011.12.008

[135] N. Lepperhoff and B. Petersdorf, "Umgang mit Datenschutzerklärungen im Internet – Ergebnisse einer empirischen Untersuchung," *Datenschutz und Datensicherheit - DuD*, vol. 33, no. 1, pp. 15–19, Jan 2009. DOI: 10.1007/s11623-009-0005-7

[136] J. Mohan, M. Wasserman, and V. Chidambaram, "Analyzing GDPR Compliance Through the Lens of Privacy Policy," in *Heterogeneous Data Management, Polystores, and Analytics for Healthcare*, V. Gadepally, T. Mattson, M. Stonebraker, F. Wang, G. Luo, Y. Laing, and A. Dubovitskaya, Eds. Cham: Springer International Publishing, 2019. DOI: 10.1007/978-3-030-33752-0_6 pp. 82–95.

[137] M. Kretschmer, J. Pennekamp, and K. Wehrle, "Cookie Banners and Privacy Policies: Measuring the Impact of the GDPR on the Web," *ACM Trans. Web*, vol. 15, no. 4, jul 2021. DOI: 10.1145/3466722

[138] N. M. Müller, D. Kowatsch, P. Debus, D. Mirdita, and K. Böttinger, "On GDPR Compliance of Companies' Privacy Policies," in *Text, Speech, and Dialogue*, K. Ekštein, Ed. Cham: Springer International Publishing, 2019. DOI: 10.1007/978-3-030-27947-9_13 pp. 151–159.

[139] European Parliament and Council, "Amendment of Privacy and Electronic Communications Directive," European Parliament and of the Council, Brüssel, Belgien, Directive Directive 2009/136/EC, 2009. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32009L0136

[140] H. Habib and L. F. Cranor, "Evaluating the Usability of Privacy Choice Mechanisms," in *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*. Boston, MA: USENIX Association, 2022, pp. 273–289.

[141] L. F. Cranor, "Cookie Monster," *Commun. ACM*, vol. 65, no. 7, p. 30–32, jun 2022. DOI: 10.1145/3538639

[142] M. Nouwens, I. Liccardi, M. Veale, D. Karger, and L. Kagal, "Dark Patterns after the GDPR: Scraping Consent Pop-Ups and Demonstrating Their Influence," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020. DOI: 10.1145/3313831.3376321 p. 1–13.

[143] N. Luhmann, *Vertrauen: Ein Mechanismus der Reduktion sozialer Komplexität*, 5th ed. München, Germany: UVK Verlagsgesellschaft mbH, 2014.

[144] D. H. McKnight and N. L. Chervany, "What is trust? A conceptual analysis and an interdisciplinary model," in *Americas Conference on Information Systems (AMCIS 2000 Proceedings)*, 2000, pp. 826–833.

[145] L. Huemer and R. J. von Krogh, Georg, "Knowledge and the Concept of Trust," in *Knowing in Firms: Understanding, Managing and Measuring Knowledge*. SAGE Publications Ltd, 1998, ch. 5, pp. 123–145. DOI: 10.4135/9781446280256.n6

[146] H. Höhmann and E. Malieva, "The concept of trust: Some notes on definitions, forms and sources," in *Trust and Entrepreneurship: A West-East Perspective*. Northampton, Massachusetts, USA: Edward Elgar Publishing Limited, 2005, pp. 7–23.

[147] ISO Central Secretary, "Systems and software engineering — Systems and software quality requirements and evaluation (SQuaRE) — Measurement of quality in use," International Organization for Standardization, Geneva, CH, Standard ISO/IEC 25022:2016, 2016. [Online]. Available: https://www.iso.org/standard/35746.html

[148] P. Giorgini, F. Massacci, J. Mylopoulos, and N. Zannone, "Requirements Engineering Meets Trust Management," in *Trust Management*, C. Jensen, S. Poslad, and T. Dimitrakos, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 176–190.

[149] G. Elahi and E. Yu, "Trust Trade-off Analysis for Security Requirements Engineering," in *2009 17th IEEE International Requirements Engineering Conference*, 2009. DOI: 10.1109/RE.2009.12 pp. 243–248.

[150] S. Nagulendra and J. Vassileva, "Providing awareness, explanation and control of personalized filtering in a social networking site," *Information Systems Frontiers*, vol. 18, no. 1, pp. 145–158, Feb 2016. DOI: 10.1007/s10796-015-9577-y

[151] T. Chakraborti, S. Sreedharan, S. Grover, and S. Kambhampati, "Plan Explanations as Model Reconciliation – An Empirical Study," in *2019 14th ACM/IEEE International*

Conference on Human-Robot Interaction (HRI), 2019. DOI: 10.1109/HRI.2019.8673193 pp. 258–266.

[152] E. S. Dahl, "Appraising Black-Boxed Technology: the Positive Prospects," *Philosophy & Technology*, vol. 31, no. 4, pp. 571–591, Dec 2018. DOI: 10.1007/s13347-017-0275-1

[153] L. Floridi, J. Cowls, M. Beltrametti, R. Chatila, P. Chazerand, V. Dignum, C. Luetge, R. Madelin, U. Pagallo, F. Rossi, B. Schafer, P. Valcke, and E. Vayena, "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations," *Minds and Machines*, vol. 28, no. 4, pp. 689–707, Dec 2018. DOI: 10.1007/s11023-018-9482-5

[154] L. Kästner, M. Langer, V. Lazar, A. Schomäcker, T. Speith, and S. Sterz, "On the Relation of Trust and Explainability: Why to Engineer for Trustworthiness," in *2021 IEEE 29th International Requirements Engineering Conference Workshops (REW)*, 2021. DOI: 10.1109/REW53955.2021.00031 pp. 169–175.

[155] T. Miller, "Explanation in Artificial Intelligence: Insights from the Social Sciences," *Artificial Intelligence*, vol. 267, pp. 1–38, 2019. DOI: 10.1016/j.artint.2018.07.007

[156] F. Doshi-Velez and B. Kim, "Towards A Rigorous Science of Interpretable Machine Learning," *arXiv preprint arXiv:1702.08608*, 2017. DOI: 10.48550/arXiv.1702.08608

[157] R. Tomsett, D. Braines, D. Harborne, A. Preece, and S. Chakraborty, "Interpretable to Whom? A Role-based Model for Analyzing Interpretable Machine Learning Systems," *arXiv preprint arXiv:1806.07552*, 2018. DOI: 10.48550/arXiv.1806.07552

[158] J. Masthoff, N. Oren, K. van Deemter, and W. W. Vasconcelos, "Towards scrutable autonomous systems," in *Symposium: Influencing People with Information*, Aberdeen, Scotland, 2012.

[159] Z. C. Lipton, "The Mythos of Model Interpretability," *Commun. ACM*, vol. 61, no. 10, p. 36–43, sep 2018. DOI: 10.1145/3233231. [Online]. Available: https://doi.org/10.1145/3233231

[160] W. Samek, T. Wiegand, and K.-R. Müller, "Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models," *arXiv preprint arXiv:1708.08296*, 2017. DOI: 10.48550/arXiv.1708.08296

[161] B. Lepri, N. Oliver, E. Letouzé, A. Pentland, and P. Vinck, "Fair, transparent, and accountable algorithmic decision-making processes," *Philosophy & Technology*, vol. 31, no. 4, pp. 611–627, 2018. DOI: 10.1007/s13347-017-0279-x

[162] S. Robbins, "A Misdirected Principle with a Catch: Explicability for AI," *Minds and Machines*, vol. 29, no. 4, pp. 495–514, 2019. DOI: 10.1007/s11023-019-09509-3

[163] M. Krishnan, "Against Interpretability: A Critical Examination of the Interpretability Problem in Machine Learning," *Philosophy & Technology*, pp. 1–16, 2019. DOI: 10.1007/s13347-019-00372-9

[164] W. Pieters, "Explanation and trust: what to tell the user in security and AI?" *Ethics and Information Technology*, vol. 13, no. 1, pp. 53–64, Mar 2011. DOI: 10.1007/s10676-010-9253-3

[165] L. Chazette and K. Schneider, "Explainability as a non-functional requirement: challenges and recommendations," *Requirements Engineering*, vol. 25, no. 4, pp. 493–514, 2020. DOI: 10.1007/s00766-020-00333-1

[166] L. Chazette, O. Karras, and K. Schneider, "Do End-Users Want Explanations? Analyzing the Role of Explainability as an Emerging Aspect of Non-Functional Requirements," in *2019 IEEE 27th International Requirements Engineering Conference (RE)*, 2019. DOI: 10.1109/re.2019.00032 pp. 223–233.

[167] C. J. Cai, J. Jongejan, and J. Holbrook, "The effects of example-based explanations in a machine learning interface," in *Proceedings of the 24th International Conference on Intelligent User Interfaces*, 2019, pp. 258–262.

[168] R. F. Kizilcec, "How Much Information? Effects of Transparency on Trust in an Algorithmic Interface," in *Proceedings of the 2016 Conference on Human Factors in Computing Systems (CHI)*. New York, NY, USA: Association for Computing Machinery, 2016. DOI: 10.1145/2858036.2858402 pp. 2390–2395.

[169] A. Springer and S. Whittaker, "Progressive disclosure: empirically motivated approaches to designing effective transparency," in *Proceedings of the 24th international conference on intelligent user interfaces*, 2019, pp. 107–120.

[170] M. Langer, D. Oster, T. Speith, H. Hermanns, L. Kästner, E. Schmidt, A. Sesing, and K. Baum, "What do we want from Explainable Artificial Intelligence (XAI)? – A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research," *Artificial Intelligence*, vol. 296, p. 103473, 2021. DOI: 10.1016/j.artint.2021.103473

[171] L. Chazette, "Mitigating Challenges in the Elicitation and Analysis of Transparency Requirements," in *2019 IEEE 27th International Requirements Engineering Conference (RE)*, 2019. DOI: 10.1109/RE.2019.00064 pp. 470–475.

[172] L. Chazette, J. Klünder, M. Balci, and K. Schneider, "How Can We Develop Explainable Systems? Insights from a Literature Review and an Interview Study," in *Proceedings of the International Conference on Software and System Processes and International Conference on Global Software Engineering*, ser. ICSSP'22. New York, NY, USA: Association for Computing Machinery, 2022. DOI: 10.1145/3529320.3529321 p. 1–12.

[173] H. Deters, J. Droste, and K. Schneider, "A Means to What End? Evaluating the Explainability of Software Systems Using Goal-Oriented Heuristics," in *Proceedings of the 27th International Conference on Evaluation and Assessment in Software Engineering*, ser. EASE '23. New York, NY, USA: Association for Computing Machinery, 2023. DOI: 10.1145/3593434.3593444 p. 329–338.

[174] K. Renaud and L. A. Shepherd, "How to Make Privacy Policies both GDPR-Compliant and Usable," in *2018 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, 2018. DOI: 10.1109/CyberSA.2018.8551442 pp. 1–8.

[175] A. M. McDonald, R. W. Reeder, P. G. Kelley, and L. F. Cranor, "A Comparative Study of Online Privacy Policies and Formats," in *Privacy Enhancing Technologies*, I. Goldberg and M. J. Atallah, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 37–55.

[176] L. Cranor, "P3P: making privacy policies more useful," *IEEE Security & Privacy*, vol. 1, no. 6, pp. 50–55, 2003. DOI: 10.1109/MSECP.2003.1253568

[177] L. F. Cranor, P. Guduru, and M. Arjula, "User Interfaces for Privacy Agents," *ACM Trans. Comput.-Hum. Interact.*, vol. 13, no. 2, p. 135–178, Jun. 2006. DOI: 10.1145/1165734.1165735

[178] S. Zimmeck and S. M. Bellovin, "Privee: An Architecture for Automatically Analyzing Web Privacy Policies," in *23rd USENIX Security Symposium (USENIX Security 14)*. San Diego, CA: USENIX Association, Aug. 2014, pp. 1–16.

[179] H. Harkous, K. Fawaz, R. Lebret, F. Schaub, K. G. Shin, and K. Aberer, "Polisis: Automated Analysis and Presentation of Privacy Policies Using Deep Learning," in *27th USENIX Security Symposium (USENIX Security 18)*. Baltimore, MD: USENIX Association, 2018, pp. 531–548.

[180] S. Zimmeck, P. Story, D. Smullen, A. Ravichander, Z. Wang, J. R. Reidenberg, N. C. Russell, and N. M. Sadeh, "MAPS: Scaling Privacy Compliance Analysis to a Million Apps," *Proceedings on Privacy Enhancing Technologies*, vol. 2019, no. 3, pp. 66–86, 2019. DOI: 10.2478/popets-2019-0037

[181] B. Andow, S. Y. Mahmud, W. Wang, J. Whitaker, W. Enck, B. Reaves, K. Singh, and T. Xie, "PolicyLint: Investigating Internal Privacy Policy Contradictions on Google Play," in *28th USENIX Security Symposium (USENIX Security 19)*. Santa Clara, CA: USENIX Association, Aug. 2019, pp. 585–602.

[182] W. B. Tesfay, P. Hofmann, T. Nakamura, S. Kiyomoto, and J. Serna, "PrivacyGuide: Towards an Implementation of the EU GDPR on Internet Privacy Policy Evaluation," in *Proceedings of the Fourth ACM International Workshop on Security and Privacy Analytics*, ser. IWSPA '18. New York, NY, USA: Association for Computing Machinery, 2018. DOI: 10.1145/3180445.3180447 p. 15–21.

[183] R. Nokhbeh Zaeem, S. Anya, A. Issa, J. Nimergood, I. Rogers, V. Shah, A. Srivastava, and K. S. Barber, "PrivacyCheck v2: A Tool That Recaps Privacy Policies for You," in *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, ser. CIKM '20. New York, NY, USA: Association for Computing Machinery, 2020. DOI: 10.1145/3340531.3417469 p. 3441–3444.

[184] F. Schaub, R. Balebako, A. L. Durity, and L. F. Cranor, "A Design Space for Effective Privacy Notices," in *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*. Ottawa: USENIX Association, 2015, pp. 1–17.

[185] J. M. Robillard, T. L. Feng, A. B. Sporn, J.-A. Lai, C. Lo, M. Ta, and R. Nadler, "Availability, Readability, and Content of Privacy Policies and Terms of Agreements of Mental Health Apps," *Internet Interventions*, vol. 17, p. 100243, 2019. DOI: 10.1016/j.invent.2019.100243

[186] A. Kitkowska, M. Warner, Y. Shulman, E. Wästlund, and L. A. Martucci, "Enhancing Privacy through the Visual Design of Privacy Notices: Exploring the Interplay of Curiosity, Control and Affect," in *Proceedings of the Sixteenth USENIX Conference on Usable Privacy and Security*, ser. Soups'20. Usa: USENIX Association, 2020.

[187] F. Schaub, B. Könings, and M. Weber, "Context-Adaptive Privacy: Leveraging Context Awareness to Support Privacy Decision Making," *IEEE Pervasive Computing*, vol. 14, no. 1, pp. 34–43, 2015. DOI: 10.1109/mprv.2015.5

[188] A.-M. Ortloff, L. Güntner, M. Windl, D. Feth, and S. Polst, "Evaluation kontextueller Datenschutzerklärungen," in *Mensch und Computer 2018 - Workshopband*, R. Dachselt and G. Weber, Eds. Bonn: Gesellschaft für Informatik e.V., 2018. DOI: 10.18420/muc2018-ws08-0541

[189] C. I. Wickramasinghe and D. Reinhardt, "A User-Centric Privacy-Preserving Approach to Control Data Collection, Storage, and Disclosure in Own Smart Home Environments,"

in *Mobile and Ubiquitous Systems: Computing, Networking and Services*, T. Hara and H. Yamaguchi, Eds. Cham: Springer International Publishing, 2022, pp. 190–206.

[190] R. Y. Wong and D. K. Mulligan, "Bringing Design to the Privacy Table: Broadening "Design" in "Privacy by Design" Through the Lens of HCI," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI '19. New York, NY, USA: Association for Computing Machinery, 2019. DOI: 10.1145/3290605.3300492 p. 1–17.

[191] O. Kulyk, P. Gerber, K. Marky, C. Beckmann, and M. Volkamer, "Does this app respect my privacy? Design and evaluation of information materials supporting privacy-related decisions of smartphone users," in *Workshop on usable security (USEC'19). San Diego, CA*, 2019, pp. 1–10.

[192] Y. Shulman, A. Kitkowska, and J. Meyer, "Informing Users: Effects of Notification Properties and User Characteristics on Sharing Attitudes," *International Journal of Human–Computer Interaction*, vol. 0, no. 0, pp. 1–29, 2022. DOI: 10.1080/10447318.2022.2086592

[193] P. G. Kelley, L. Cesca, J. Bresee, and L. F. Cranor, "Standardizing Privacy Notices: An Online Study of the Nutrition Label Approach," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. Chi '10. New York, NY, USA: Association for Computing Machinery, 2010. DOI: 10.1145/1753326.1753561 p. 1573–1582.

[194] P. G. Kelley, J. Bresee, L. F. Cranor, and R. W. Reeder, "A Nutrition Labelfor Privacy," in *Proceedings of the 5th Symposium on Usable Privacy and Security*, ser. SOUPS '09. New York, NY, USA: Association for Computing Machinery, 2009. DOI: 10.1145/1572532.1572538

[195] P. Benassi, "TRUSTe: An Online Privacy Seal Program," *Commun. ACM*, vol. 42, no. 2, p. 56–59, feb 1999. DOI: 10.1145/293411.293461

[196] R. Rodrigues, D. Wright, and K. Wadhwa, "Developing a privacy seal scheme (that works)," *International Data Privacy Law*, vol. 3, no. 2, pp. 100–116, 02 2013. DOI: 10.1093/idpl/ips037

[197] T. Moores, "Do Consumers Understand the Role of Privacy Seals in E-Commerce?" *Commun. ACM*, vol. 48, no. 3, p. 86–91, mar 2005. DOI: 10.1145/1047671.1047674. [Online]. Available: https://doi.org/10.1145/1047671.1047674

[198] M. B. Hosseini, J. Heaps, R. Slavin, J. Niu, and T. Breaux, "Ambiguity and Generality in Natural Language Privacy Policies," in *2021 IEEE 29th International Requirements Engineering Conference (RE)*, 2021. DOI: 10.1109/RE51729.2021.00014 pp. 70–81.

[199] T. D. Breaux, H. Hibshi, and A. Rao, "Eddy, a formal language for specifying and analyzing data flow specifications for conflicting privacy requirements," *Requirements Engineering*, vol. 19, no. 3, pp. 281–307, Sep 2014. DOI: 10.1007/s00766-013-0190-7

[200] A. K. Massey, J. Eisenstein, A. I. Antón, and P. P. Swire, "Automated text mining for requirements analysis of policy documents," in *2013 21st IEEE International Requirements Engineering Conference (RE)*, 2013. DOI: 10.1109/RE.2013.6636700 pp. 4–13.

[201] A. K. Massey, E. Holtgrefe, and S. Ghanavati, "Modeling Regulatory Ambiguities for Requirements Analysis," in *Conceptual Modeling*, H. C. Mayr, G. Guizzardi, H. Ma, and O. Pastor, Eds. Cham: Springer International Publishing, 2017, pp. 231–238.

[202] E. Kamsties and B. Paech, "Taming Ambiguity in Natural Language Requirements," in *International Conference Software and Systems Engineering and their Applications (ICSSAE) 2000*, Paris, France, 2000.

[203] A. R. Hevner, "A three cycle view of design science research," *Scandinavian journal of information systems*, vol. 19, no. 2, p. 7, 2007.

[204] J. F. Nunamaker Jr., C. Minder, and . P. Titus D.M, "Systems Development in Information Systems Research," *Journal of Management Information Systems*, vol. 7, no. 3, pp. 89–106, 1990. DOI: 10.1080/07421222.1990.11517898

[205] J. Iivari and J. Venable, "Action research and design science research - Seemingly similar but decisively dissimilar," in *17th European Conference on Information Systems, ECIS 2009, Verona, Italy, 2009*, S. Newell, E. A. Whitley, N. Pouloudi, J. Wareham, and L. Mathiassen, Eds., 2009, pp. 1642–1653.

[206] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A Design Science Research Methodology for Information Systems Research," *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45–77, 2007. DOI: 10.2753/mis0742-1222240302

[207] H. A. Simon, *The Sciences of the Artificial*, 3rd ed. Cambridge, MA, USA: The MIT Press, 1996.

[208] V. Wilson, "Research Methods: Triangulation," *Evidence Based Library and Information Practice*, vol. 11, no. 1(s), p. 66–68, Mar. 2016. DOI: 10.18438/b86s5f

[209] U. Flick, *The Sciences An Introduction to Qualitative Research*, 4th ed. Thousand Oaks, CA, USA: SAGE Publications Ltd, 2009.

[210] W. R. Shadish, T. D. Cook, and D. T. Campbell, *Experimental and quasi-experimental designs for generalized causal inference*, ser. Experimental and quasi-experimental designs for generalized causal inference. Boston, MA, US: Houghton, Mifflin and Company, 2002.

[211] D. I. K. Sjøberg, T. Dybå, B. C. D. Anda, and J. E. Hannay, "Building Theories in Software Engineering," in *Guide to Advanced Empirical Software Engineering*, F. Shull, J. Singer, and D. I. K. Sjøberg, Eds. London: Springer London, 2008, pp. 312–336. DOI: 10.1007/978-1-84800-044-5_12

[212] B. G. Buchanan and E. H. Shortliffe, *Rule-based expert systems: the MYCIN experiments of the Stanford Heuristic Programming Project.* Addison-Wesley, 1984.

[213] F. Kratzert, M. Herrnegger, D. Klotz, S. Hochreiter, and G. Klambauer, "Do internals of neural networks make sense in the context of hydrology?" in *AGU Fall Meeting Abstracts*, vol. 2018, 2018, pp. H13B–06.

[214] A. Adadi and M. Berrada, "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," *IEEE Access*, vol. 6, pp. 52 138–52 160, 2018. DOI: 10.1109/access.2018.2870052

[215] J. P. Winkler and A. Vogelsang, ""What Does My Classifier Learn?Ä Visual Approach to Understanding Natural Language Text Classifiers," in *Natural Language and Information Systems*, F. Frasincar, A. Ittoo, L. M. Nguyen, and E. Métais, Eds., 2017. DOI: 10.1007/978-3-319-59569-6_55 pp. 468–479.

[216] A. Bussone, S. Stumpf, and D. O'Sullivan, "The Role of Explanations on Trust and Reliance in Clinical Decision Support Systems," in *2015 International Conference on Healthcare Informatics.* New York, NY, USA: Ieee, 2015. DOI: 10.1109/ichi.2015.26 pp. 160–169.

[217] M. Hind, D. Wei, M. Campbell, N. C. F. Codella, A. Dhurandhar, A. Mojsilović, K. Natesan Ramamurthy, and K. R. Varshney, "TED: Teaching AI to Explain Its Decisions," in *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society.* New York, NY, USA: Acm, 2019. DOI: 10.1145/3306618.3314273 pp. 123–129.

[218] A. Rosenfeld and A. Richardson, "Explainability in human-agent systems," *Autonomous Agents and Multi-Agent Systems*, vol. 33, no. 6, pp. 673–705, 2019. DOI: 10.1007/s10458-019-09408-y

[219] D. Mairiza and D. Zowghi, "Constructing a Catalogue of Conflicts among Non-functional Requirements," in *Evaluation of Novel Approaches to Software Engineering*. Berlin/Heidelberg, DE: Springer, 2011, pp. 31–44.

[220] F.-L. Li, J. Horkoff, J. Mylopoulos, R. S. S. Guizzardi, G. Guizzardi, A. Borgida, and L. Liu, "Non-functional requirements as qualities, with a spice of ontology," in *2014 IEEE 22nd International Requirements Engineering Conference (RE)*, 2014. DOI: 10.1109/re.2014.6912271 pp. 293–302.

[221] A. Hoffmann, E. A. C. Bittner, and J. M. Leimeister, "The Emergence of Mutual and Shared Understanding in the System Development Process," in *Requirements Engineering: Foundation for Software Quality*, J. Doerr and A. L. Opdahl, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 174–189.

[222] T. Margaret, "Establishing Mutual Understanding in Systems Design: An Empirical Study," *Journal of Management Information Systems*, vol. 10, no. 4, pp. 159–182, 1994. DOI: 10.1080/07421222.1994.11518024

[223] G. J. De Vreede, R. O. Briggs, and A. P. Massey, "Collaboration engineering: foundations and opportunities: editorial to the special issue on the journal of the association of information systems," *Journal of the Association for Information Systems*, vol. 10, no. 3, p. 7, 2009. DOI: 10.17705/1jais.00191

[224] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering," Keele University, Tech. Rep., 2007.

[225] C. Wohlin, "Guidelines for snowballing in systematic literature studies and a replication in software engineering," in *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, New York, NY, USA, 2014. DOI: 10.1145/2601248.2601268 pp. 1–10.

[226] H. Zhang and M. Ali Babar, "On Searching Relevant Studies in Software Engineering," in *Proceedings of the 14th International Conference on Evaluation and Assessment in Software Engineering*, ser. Ease'10. Swindon, GBR: BCS Learning & Development Ltd., 2010, p. 111–120.

[227] J. L. Fleiss, "Measuring nominal scale agreement among many raters," *Psychological Bulletin*, vol. 76, no. 5, pp. 378–382, 1971. DOI: 10.1037/h0031619

[228] J. R. Landis and G. G. Koch, "The Measurement of Observer Agreement for Categorical Data," *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977. DOI: 10.2307/2529310

[229] J. F. Wolfswinkel, E. Furtmueller, and C. P. M. Wilderom, "Using grounded theory as a method for rigorously reviewing literature," *European journal of Information Systems*, vol. 22, no. 1, pp. 45–55, 2013. DOI: 10.1057/ejis.2011.51

[230] J. Saldaña, *The Coding Manual for Qualitative Researchers*, 2nd ed. Thousand Oaks, CA, USA: SAGE Publications Inc., 2013.

[231] K. Charmaz, *Constructing grounded theory: A practical guide through qualitative analysis*, 2nd ed. Thousand Oaks, CA, USA: SAGE Publications Inc., 2014.

[232] M. B. Miles and A. M. Huberman, *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks, CA, USA: SAGE Publications, 1994.

[233] R. E. Boyatzis, *Transforming Qualitative Information: Thematic Analysis and Code Development*. Thousand Oaks, CA, USA: SAGE Publications, 1998.

[234] A. Yasin, R. Fatima, L. Wen, W. Afzal, M. Azhar, and R. Torkar, "On Using Grey Literature and Google Scholar in Systematic Literature Reviews in Software Engineering," *IEEE Access*, vol. 8, pp. 36 226–36 243, 2020. DOI: 10.1109/access.2020.2971712

[235] L. Cranor, M. Langheinrich, M. Marchiori, M. Presler-Marshall, and J. Reagle, "The Platform for Privacy Preferences 1.0 (P3P1.0) Specification," https://www.w3.org/TR/P3P/, 2002, Letzter Zugriff: 27.02.2023.

[236] L. Chung, B. A. Nixon, E. Yu, and J. Mylopoulos, *Non-Functional Requirements in Software Engineering*. Boston, MA, USA: Springer Science & Business Media, 2012.

[237] D. V. Carvalho, E. M. Pereira, and J. S. Cardoso, "Machine learning interpretability: A survey on methods and metrics," *Electronics*, vol. 8, no. 8, 2019. DOI: 10.3390/electronics8080832

[238] A. Holzinger, G. Langs, H. Denk, K. Zatloukal, and H. Müller, "Causability and explainability of artificial intelligence in medicine," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 9, no. 4, pp. 1–13, 2019. DOI: 10.1002/widm.1312

[239] J. Hois, D. Theofanou-Fuelbier, and A. J. Junk, "How to Achieve Explainability and Transparency in Human AI Interaction," in *International Conference on Human-Computer Interaction (HCI)*. Cham, CH: Springer International Publishing, 2019. DOI: 10.1007/978-3-030-23528-4_25 pp. 177–183.

[240] A. Glass, D. L. McGuinness, and M. Wolverton, "Toward establishing trust in adaptive agents," in *Proceedings of the 13th International Conference on Intelligent User Interfaces (IUI)*. New York, NY, USA: Acm, 2008. DOI: 10.1145/1378773.1378804 pp. 227–236.

[241] Q. V. Liao, D. M. Gruen, and S. Miller, "Questioning the AI: Informing Design Practices for Explainable AI User Experiences," in *Proceedings of the 2020 Conference on*

Human Factors in Computing Systems (CHI). New York, NY, USA: Acm, 2020. DOI: 10.1145/3313831.3376590 pp. 1–15.

[242] P. Dourish, "What we talk about when we talk about context," *Personal and Ubiquitous Computing*, vol. 8, no. 1, pp. 19–30, 2004. DOI: 10.1007/s00779-003-0253-8

[243] M. Langer, D. Oster, T. Speith, H. Hermanns, L. Kästner, E. Schmidt, A. Sesing, and K. Baum, "What Do We Want From Explainable Artificial Intelligence (XAI)? – A Stakeholder Perspective on XAI and a Conceptual Model Guiding Interdisciplinary XAI Research," *Artificial Intelligence*, 2021. DOI: 10.1016/j.artint.2021.103473

[244] M. T. Ribeiro, S. Singh, and C. Guestrin, ""Why Should I Trust You?": Explaining the Predictions of Any Classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: Acm, 2016. DOI: 10.1145/2939672.2939778 pp. 1135–1144.

[245] A. B. Arrieta, N. Díaz-Rodríguez, J. D. Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Information Fusion*, vol. 58, pp. 82–115, 2020. DOI: 10.1016/j.inffus.2019.12.012

[246] I. Lage, D. Lifschitz, F. Doshi-Velez, and O. Amir, "Exploring Computational User Models for Agent Policy Summarization," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI)*, 2019, pp. 59–65.

[247] J. McInerney, B. Lacker, S. Hansen, K. Higley, H. Bouchard, A. Gruson, and R. Mehrotra, "Explore, exploit, and explain: personalizing explainable recommendations with bandits," in *Proceedings of the 12th ACM Conference on Recommender Systems (RecSys)*. New York, NY, USA: Acm, 2018. DOI: 10.1145/3240323.3240354 pp. 31–39.

[248] G. Friedrich and M. Zanker, "A taxonomy for generating explanations in recommender systems," *AI Magazine*, vol. 32, no. 3, pp. 90–98, 2011.

[249] T. Eiter, Z. G. Saribatur, and P. Schüller, "Abstraction for Zooming-In to Unsolvability Reasons of Grid-Cell Problems," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 7–13.

[250] R. Binns, M. Van Kleek, M. Veale, U. Lyngs, J. Zhao, and N. Shadbolt, "'It's Reducing a Human Being to a Percentage': Perceptions of Justice in Algorithmic Decisions," in *Proceedings of the 2018 Conference on Human Factors in Computing Systems (CHI)*. New York, NY, USA: Acm, 2018. DOI: 10.1145/3173574.3173951 pp. 1–14.

[251] Z. Juozapaitis, A. Koul, A. Fern, M. Erwig, and F. Doshi-Velez, "Explainable Reinforcement Learning via Reward Decomposition," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 47–53.

[252] I. Nunes and D. Jannach, "A systematic review and taxonomy of explanations in decision support and recommender systems," *User Modeling and User-Adapted Interaction*, vol. 27, no. 3-5, pp. 393–444, 2017. DOI: 10.1007/s11257-017-9195-0

[253] E. S. Vorm, "Assessing Demand for Transparency in Intelligent Systems Using Machine Learning," in *2018 Innovations in Intelligent Systems and Applications (INISTA)*. Ieee, 2018. DOI: 10.1109/inista.2018.8466328 pp. 1–7.

[254] H. Cramer, V. Evers, S. Ramlal, V. S. Maarten, L. Rutledge, N. Stash, L. Aroyo, and B. Wielinga, "The effects of transparency on trust in and acceptance of a content-based art recommender," *User Modeling and User-adapted interaction*, vol. 18, no. 5, p. 455, 2008. DOI: 10.1007/s11257-008-9051-3

[255] A. Finkelstein and J. Dowell, "A comedy of errors: the London Ambulance Service case study," in *Proceedings of the 8th International Workshop on Software Specification and Design*, 1996. DOI: 10.1109/iwssd.1996.501141 pp. 2–4.

[256] C. Cappelli, H. Cunha, B. Gonzalez-Baixauli, and J. C. S. do Prado Leite, "Transparency versus Security: Early Analysis of Antagonistic Requirements," in *Proceedings of the 2010 ACM Symposium on Applied Computing*, ser. Sac '10. New York, NY, USA: Association for Computing Machinery, 2010. DOI: 10.1145/1774088.1774151 p. 298–305.

[257] K. Sokol and P. A. Flach, "Explainability fact sheets: A framework for systematic assessment of explainable approaches," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. New York, NY, USA: Association for Computing Machinery, 2020. DOI: 10.1145/3351095.3372870 pp. 56–67.

[258] K. Sokol and P. Flach, "One Explanation Does Not Fit All," *KI-Künstliche Intelligenz*, vol. 34, no. 2, pp. 235–250, 2020. DOI: 10.1007/s13218-020-00637-y

[259] 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)*. Ieee, 2018, pp. 382–387.

[260] J. Schneider and J. P. Handali, "Personalized Explanation For Machine Learning: A Conceptualization," in *27th European Conference on Information Systems (ECIS)*, 2019.

[261] C. Henin and L. M. Daniel, "Towards a Generic Framework for Black-Box Explanation Methods," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI)*, 2019, pp. 28–34.

[262] J. Gluck, F. Schaub, A. Friedman, H. Habib, N. Sadeh, L. F. Cranor, and Y. Agarwal, "How Short is Too Short? Implications of Length and Framing on the Effectiveness of Privacy Notices," in *Proceedings of the Twelfth USENIX Conference on Usable Privacy and Security*, ser. Soups '16. Usa: USENIX Association, 2016, p. 321–340.

[263] J. Earp, A. Anton, and O. Jarvinen, "A Social, Technical, and Legal Framework for Privacy Management and Policies," in *Eighth Americas Conference on Information Systems (AMCIS 2002 Proceedings)*. Association for Information Systems, 2002, pp. 605–612.

[264] L. Chen, D. Yan, and F. Wang, "User Evaluations on Sentiment-Based Recommendation Explanations," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 9, no. 4, pp. 1–38, 2019. DOI: 10.1145/3282878

[265] J. Bhatia, M. C. Evans, and T. D. Breaux, "Identifying incompleteness in privacy policy goals using semantic frames," *Requirements Engineering*, vol. 24, no. 3, pp. 291–313, Sep 2019. DOI: 10.1007/s00766-019-00315-y

[266] M. Barhamgi, C. Perera, C. Ghedira, and D. Benslimane, "User-centric Privacy Engineering for the Internet of Things," *IEEE Cloud Computing*, vol. 5, no. 5, pp. 47–57, 2018. DOI: 10.1109/mcc.2018.053711666

[267] S. Samat and A. Acquisti, "Format vs. Content: The Impact of Risk and Presentation on Disclosure Decisions," in *Proceedings of the Thirteenth USENIX Conference on Usable Privacy and Security*, ser. Soups '17. Usa: USENIX Association, 2017, p. 377–384.

[268] R. Sloan, Robert H. Warner, "Beyond Notice and Choice: Privacy, Norms, and Consent," *Journal of High Technology Law*, vol. 14, p. 370, 2014.

[269] T. D. Huynh, N. Tsakalakis, A. Helal, S. Stalla-Bourdillon, and L. Moreau, *Explainability-by-Design: A Methodology to Support Explanations in Decision-Making Systems*. arXiv, 2022.

[270] S. Killingsworth, "Minding Your own Business: Privacy Policies in Principle and in Practice," *Journal of Intellectual Property Law,*, vol. 7, no. 1, pp. 39–88, 1999.

[271] C. Flavián and M. Guinalíu, "Consumer trust, perceived security and privacy policy," *Industrial Management & Data Systems*, vol. 106, no. 5, pp. 601–620, Jan 2006. DOI: 10.1108/02635570610666403

[272] J. Angulo and M. Ortlieb, "„WTH..!?!" Experiences, Reactions, and Expectations Related to Online Privacy Panic Situations," in *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*. Ottawa: USENIX Association, 2015, pp. 19–38.

[273] M. Corrales Compagnucci, M. Fenwick, H. Haapio, T. Minssen, and E. P. Vermeulen, "Technology-Driven Disruption of Healthcare &'UI Layer'Privacy-by-Design," *Cambridge Bioethics and the Law series, Cambridge University Press, Forthcoming*, 2020. DOI: 10.2139/ssrn.3611702

[274] M. North, "An Examination of Mobile App Privacy Policies and Thirdparty Data Sharing," *Issues in Information Systems*, vol. 14, no. 2, 2013.

[275] S. Sharma, *Data privacy and GDPR handbook*, 1st ed. John Wiley & Sons, 2019.

[276] M. Fernandes, A. R. Silva, and A. Gonçalves, "Specification of Personal Data Protection Requirements - Analysis of Legal Requirements from the GDPR Regulation," in *Proceedings of the 20th International Conference on Enterprise Information Systems, ICEIS 2018, Funchal, Madeira, Portugal, March 21-24, 2018, Volume 2*, S. Hammoudi, M. Smialek, O. Camp, and J. Filipe, Eds. SciTePress, 2018. DOI: 10.5220/0006810603980405 pp. 398–405.

[277] J. Schoenherr, "Whose Privacy, What Surveillance? Dimensions of the Mental Models for Privacy and Security," *IEEE Technology and Society Magazine*, vol. 41, no. 1, pp. 54–65, 2022. DOI: 10.1109/mts.2022.3147536

[278] F. Thoma, "How Siemens Assesses Privacy Impacts," in *Privacy Impact Assessment*, D. Wright and P. De Hert, Eds. Dordrecht: Springer Netherlands, 2012, pp. 275–284. DOI: 10.1007/978-94-007-2543-0_12

[279] L. F. Cranor, "Necessary but Not Sufficient: Standardized Mechanisms for Privacy Notice and Choice The Economics of Privacy," *Journal on Telecommunications and High Technology Law*, vol. 10, p. 273, 2012.

[280] N. Notario, A. Crespo, Y.-S. Martin, J. M. Del Alamo, D. L. Metayer, T. Antignac, A. Kung, I. Kroener, and D. Wright, "PRIPARE: Integrating Privacy Best Practices into a Privacy Engineering Methodology," in *2015 IEEE Security and Privacy Workshops*, 2015. DOI: 10.1109/spw.2015.22 pp. 151–158.

[281] L. Zhang-Kennedy and S. Chiasson, "„Whether it's moral is a whole other story": Consumer perspectives on privacy regulations and corporate data practices," in *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. USENIX Association, 2021, pp. 197–216.

[282] S. Spiekermann and L. F. Cranor, "Engineering Privacy," *IEEE Transactions on Software Engineering*, vol. 35, no. 1, pp. 67–82, 2009. DOI: 10.1109/tse.2008.88

[283] A. Antón, J. Earp, and A. Reese, "Analyzing Website privacy requirements using a privacy goal taxonomy," in *Proceedings IEEE Joint International Conference on Requirements Engineering*, 2002. DOI: 10.1109/icre.2002.1048502 pp. 23–31.

[284] H. Peschke, *Betroffenenorientierte Systementwicklung: Prozess Und Methoden Der Entwicklung Menschengerechter Informationssysteme.* Frankfurt am Main, Germany: Peter Lang GmbH, 1986.

[285] P. A. Norberg and D. R. Horne, "Privacy Attitudes and Privacy-Related Behavior," *Psychology & Marketing*, vol. 24, no. 10, pp. 829–847, 2007. DOI: 10.1002/mar.20186

[286] T. Dienlin and S. Trepte, "Is the privacy paradox a relic of the past? An in-depth analysis of privacy attitudes and privacy behaviors," *European Journal of Social Psychology*, vol. 45, no. 3, pp. 285–297, 2015. DOI: 10.1002/ejsp.2049

[287] P. B. Brandtzaeg, A. Pultier, and G. M. Moen, "Losing control to data-hungry apps: A mixed-methods approach to mobile app privacy," *Social Science Computer Review*, vol. 37, no. 4, pp. 466–488, 2019. DOI: 10.1177/0894439318777706

[288] Y. Guo, X. Wang, and C. Wang, "Impact of privacy policy content on perceived effectiveness of privacy policy: the role of vulnerability, benevolence and privacy concern," *Journal of Enterprise Information Management*, vol. 35, no. 3, pp. 774–795, Jan 2022. DOI: 10.1108/jeim-12-2020-0481

[289] V. M. Wottrich, E. A. van Reijmersdal, and E. G. Smit, "App Users Unwittingly in the Spotlight: A Model of Privacy Protection in Mobile Apps," *Journal of Consumer Affairs*, vol. 53, no. 3, pp. 1056–1083, 2019. DOI: 10.1111/joca.12218

[290] E. Okoyomon, N. Samarin, P. Wijesekera, A. Elazari Bar On, N. Vallina-Rodriguez, I. Reyes, Á. Feal, S. Egelman *et al.*, "On the ridiculousness of notice and consent: Contradictions in app privacy policies," in *Workshop on Technology and Consumer Protection (ConPro 2019), in conjunction with the 39th IEEE Symposium on Security and Privacy*, 2019.

[291] T. Li, E. B. Neundorfer, Y. Agarwal, and J. I. Hong, "Honeysuckle: Annotation-Guided Code Generation of In-App Privacy Notices," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 3, 2021. DOI: 10.1145/3478097

[292] S. Thomsen, "Corporate values and corporate governance," *Corporate Governance: The international journal of business in society*, vol. 4, no. 4, pp. 29–46, Jan 2004. DOI: 10.1108/14720700410558862

[293] L. Verderame, D. Caputo, A. Romdhana, and A. Merlo, "On the (Un)Reliability of Privacy Policies in Android Apps," in *2020 International Joint Conference on Neural Networks (IJCNN)*, 2020. DOI: 10.1109/ijcnn48605.2020.9206660 pp. 1–9.

[294] I. F. Alexander, "A Better Fit - Characterising the Stakeholders," in *CAiSE'04 Workshops in connection with The 16th Conference on Advanced Information Systems Engineering, Riga, Latvia, 7-11 June, 2004, Knowledge and Model Driven Information Systems Engineering for Networked Organisations, Proceedings, Vol. 2*, J. Grundspenkis and M. Kirikova, Eds. Riga, Latvia: Faculty of Computer Science and Information Technology, Riga Technical University, 2004, pp. 215–223.

[295] J. P. Carvallo, X. Franch, and C. Quer, "Managing Non-Technical Requirements in COTS Components Selection," in *14th IEEE International Requirements Engineering Conference (RE)*. New York, NY, USA: Ieee, 2006. DOI: 10.1109/re.2006.40 pp. 323–326.

[296] F. Wolfe Sharp and P. R. Sharp, "What Do You Mean My Website Isn't Accessible? Why Web Accessibility Matters in the Digital World," in *Exploring Ethical Problems in Today's Technological World*. Hershey, PA, USA: IGI Global, 2022, pp. 165–182. DOI: 10.4018/978-1-6684-5892-1.ch009

[297] S. Veys, D. Serrano, M. Stamos, M. Herman, N. Reitinger, M. L. Mazurek, and B. Ur, "Pursuing Usable and Useful Data Downloads Under GDPR/CCPA Access Rights via Co-Design," in *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. USENIX Association, 2021, pp. 217–242.

[298] R. Lämmel and E. Pek, "Understanding privacy policies," *Empirical Software Engineering*, vol. 18, no. 2, pp. 310–374, Apr 2013. DOI: 10.1007/s10664-012-9204-1

[299] A. I. Antón and J. B. Earp, "A requirements taxonomy for reducing Web site privacy vulnerabilities," *Requirements Engineering*, vol. 9, no. 3, pp. 169–185, Aug 2004. DOI: 10.1007/s00766-003-0183-z

[300] H. Harkous, K. Fawaz, K. G. Shin, and K. Aberer, "PriBots: Conversational Privacy with Chatbots," in *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*. Denver, CO: USENIX Association, 2016.

[301] T. Berners-Lee, R. T. Fielding, and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax," Internet Requests for Comments, RFC Editor, STD 66, January 2005, Letzter Zugriff: 07.03.2023. [Online]. Available: https://www.rfc-editor.org/rfc/rfc3986.txt

[302] S. Wilson, F. Schaub, A. A. Dara, F. Liu, S. Cherivirala, P. Giovanni Leon, M. Schaarup Andersen, S. Zimmeck, K. M. Sathyendra, N. C. Russell, T. B. Norton, E. Hovy, J. Reidenberg, and N. Sadeh, "The Creation and Analysis of a Website Privacy Policy Corpus," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Berlin, Germany: Association for Computational Linguistics, 2016. DOI: 10.18653/v1/P16-1126 pp. 1330–1340.

[303] A. Zolnierek and B. Rubacha, "The Empirical Study of the Naive Bayes Classifier in the Case of Markov Chain Recognition Task," in *Computer Recognition Systems, Proceedings of the 4th International Conference on Computer Recognition Systems, CORES'05, May 22-25, 2005, Rydzyna Castle, Poland*, ser. Advances in Soft Computing, M. Kurzynski, E. Puchala, M. Wozniak, and A. Zolnierek, Eds., vol. 30. Springer, 2005. DOI: 10.1007/3-540-32390-2_38 pp. 329–336.

[304] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct 2001. DOI: 10.1023/A:1010933404324

[305] D. A. Hull, "Stemming algorithms: A case study for detailed evaluation," *Journal of the American Society for Information Science*, vol. 47, no. 1, pp. 70–84, 1996. DOI: 10.1002/(SICI)1097-4571(199601)47:1<70::AID-ASI7>3.0.CO;2-#

[306] A. Kulkarni and A. Shivananda, "Converting Text to Features," in *Natural Language Processing Recipes: Unlocking Text Data with Machine Learning and Deep Learning Using Python*. Berkeley, CA: Apress, 2021, pp. 63–106. DOI: 10.1007/978-1-4842-7351-7_3

[307] S. Robertson, "Understanding inverse document frequency: on theoretical arguments for IDF," *Journal of Documentation*, vol. 60, no. 5, pp. 503–520, Jan 2004. DOI: 10.1108/00220410410560582

[308] P. Barr, J. Noble, and R. Biddle, "Icons R Icons," in *Proceedings of the Fourth Australasian User Interface Conference on User Interfaces 2003 - Volume 18*, ser. AUIC '03. AUS: Australian Computer Society, Inc., 2003, p. 25–32.

[309] I. Apple Computer, *Macintosh Human Interface Guidelines*. USA: Addison-Wesley Publishing Company, 1992.

[310] P. B. Andersen, "What Semiotics can and cannot do for HCI," *Knowledge-Based Systems*, vol. 14, no. 8, pp. 419–424, 2001, semiotic Approaches to User Interface Design.

[311] J. Lindley, H. A. Akmal, F. Pilling, and P. Coulton, "Researching AI Legibility through Design," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–13.

[312] F. Thouvenin, M. Glatthaar, J. Hotz, C. Ettlinger, and M. Tschudin, "Privacy Icons: Transparent auf einen Blick," *Jusletter*, no. 30.11.2020, p. online, November 2020. DOI: 10.5167/uzh-193220

[313] P. W. Jordan, *An Introduction to Usability*, 1st ed. Boca Raton, FL, USA: CRC Press, 2002.

[314] H. R. Hartson, J. C. Castillo, J. Kelso, and W. C. Neale, "Remote Evaluation: The Network as an Extension of the Usability Laboratory," in *Proceedings of the SIGCHI conference on human factors in computing systems*, ser. CHI 1996. New York, NY, USA: Association for Computing Machinery, 1996, pp. 228–235.

[315] A. B. Brush, M. Ames, and J. Davis, "A Comparison of Synchronous Remote and Local Usability Studies for an Expert Interface," in *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '04. New York, NY, USA: Association for Computing Machinery, 2004, p. 1179–1182.

[316] N. M. Bradburn, S. Sudman, and B. Wansink, *Asking Questions: The Definitive Guide to Questionnaire Design–for Market Research, Political Polls, and Social and Health Questionnaires*, 1st ed. San Francisco, CA, USA: John Wiley & Sons, 2004.

[317] R. Jacob, A. Heinz, and J. P. Décieux, *Umfrage: Einführung in die Methoden der Umfrageforschung*, 3rd ed. München, Deutschland: Oldenbourg Wissenschaftsverlag GmbH, 2013.

[318] J. Rubin and D. Chisnell, *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*, 2nd ed. Indianapolis, IN, USA: John Wiley & Sons, 2008.

[319] T. Boren and J. Ramey, "Thinking aloud: reconciling theory and practice," *IEEE Transactions on Professional Communication*, vol. 43, no. 3, pp. 261–278, 2000.

[320] J. Brooke, "SUS: A quick and dirty usability scale," in *Usability Evaluation In Industry (1st ed.)*, 1st ed., P. W. Jordan, B. Thomas, I. L. McClelland, and B. Weerdmeester, Eds. London, GB: Taylor & Francis, 1996, pp. 189–194.

[321] S. C. Peres, T. Pham, and R. Phillips, "Validation of the System Usability Scale (SUS): SUS in the Wild," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 57, no. 1, pp. 192–196, 2013.

[322] J. Sauro and J. R. Lewis, "Average Task Times in Usability Tests: What to Report?" in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '10. New York, NY, USA: Association for Computing Machinery, 2010. DOI: 10.1145/1753326.1753679 p. 2347–2350.

[323] J. Sauro and J. R. Lewis, *Quantifying the User Experience: Practical Statistics for User Research*, 1st ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2012.

[324] M. R. Islam, "Sample size and its role in Central Limit Theorem (CLT)," *Computational and Applied Mathematics Journal*, vol. 4, no. 1, pp. 1–7, 2018.

[325] A. Bangor, P. Kortum, and J. Miller, "Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale," *Journal of User Experience*, vol. 4, no. 3, p. 114–123, May 2009.

[326] L. F. Cranor, J. Reagle, and M. S. Ackerman, "Beyond concern: Understanding Net users attitudes about online privacy. AT&T Labs-Research Technical Report TR 99.4. 3," *Retrieved April*, vol. 14, p. 1999, 1999.

[327] H. Taylor, "Most people are „privacy pragmatists" who, while concerned about privacy, will sometimes trade it off for other benefits," *The Harris Poll*, vol. 17, no. 19, p. 44, 2003.

[328] L. Alkhariji, N. Alhirabi, M. N. Alraja, M. Barhamgi, O. Rana, and C. Perera, "Synthesising Privacy by Design Knowledge Toward Explainable Internet of Things Application Designing in Healthcare," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 17, no. 2s, jun 2021. DOI: 10.1145/3434186

[329] R. Hes and J. Borking, "Privacy-Enhancing Technologies: The Path to Anonymity," *(Revised Edition) Den Haag: Registratiekamer*, vol. 4, 1995.

[330] A. Cavoukian, J. Stoddart, A. Dix, I. Nemec, V. Peep, and M. Shroff, "Resolution on privacy by design," in *32nd International Conference of Data Protection and Privacy Commissioners*, Jerusalem, Israel, 2010, pp. 27–29.

[331] S. Gürses, C. Troncoso, and C. Diaz, "Engineering Privacy by Design," *Computers, Privacy & Data Protection*, vol. 14, no. 3, p. 25, 2011.

[332] J. van Rest, D. Boonstra, M. Everts, M. van Rijn, and R. van Paassen, "Designing Privacy-by-Design," in *Privacy Technologies and Policy*, B. Preneel and D. Ikonomou, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 55–72.

[333] X. Jiang, J. I. Hong, and J. A. Landay, "Approximate Information Flows: Socially-Based Modeling of Privacy in Ubiquitous Computing," in *UbiComp 2002: Ubiquitous Computing*, G. Borriello and L. E. Holmquist, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 176–193.

[334] A. Hazeyama, "Proposal of a Privacy Knowledge Base for Supporting Development of Privacy Friendly Software," *Procedia Computer Science*, vol. 176, pp. 1440–1448, 2020. DOI: 10.1016/j.procs.2020.09.154 Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 24th International Conference KES2020.

[335] J. Lenhard, L. Fritsch, and S. Herold, "A Literature Study on Privacy Patterns Research," in *2017 43rd Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, 2017. DOI: 10.1109/SEAA.2017.28 pp. 194–201.

[336] C. Thiel and M. Broy, "Privacy by Design als Win-win-Strategie für Wirtschaft und Verbraucher*innen," in *ZD.B Digital Dialogue*. Bayerisches Staasministerium für Umwelt und Verbraucherschutz, 2019, pp. 1–15.

[337] A. Senarath, M. Grobler, and N. A. G. Arachchilage, "Will They Use It or Not? Investigating Software Developers' Intention to Follow Privacy Engineering Methodologies," *ACM Trans. Priv. Secur.*, vol. 22, no. 4, nov 2019. DOI: 10.1145/3364224

[338] I. Hadar, T. Hasson, O. Ayalon, E. Toch, M. Birnhack, S. Sherman, and A. Balissa, "Privacy by designers: software developers' privacy mindset," *Empirical Software Engineering*, vol. 23, no. 1, pp. 259–289, Feb 2018. DOI: 10.1007/s10664-017-9517-1

[339] M. Peixoto, D. Ferreira, M. Cavalcanti, C. Silva, J. Vilela, J. Araújo, and T. Gorschek, "The perspective of Brazilian software developers on data privacy," *Journal of Systems and Software*, vol. 195, p. 111523, 2023. DOI: 10.1016/j.jss.2022.111523

[340] J. A. Konstan and J. Riedl, "Recommender systems: from algorithms to user experience," *User modeling and user-adapted interaction*, vol. 22, no. 1-2, pp. 101–123, 2012.

[341] A. Bunt, M. Lount, and C. Lauzon, "Are explanations always important?: a study of deployed, low-cost intelligent interactive systems," in *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces*. ACM, 2012, pp. 169–178.

[342] A. Papadimitriou, P. Symeonidis, and Y. Manolopoulos, "A generalized taxonomy of explanations styles for traditional and social recommender systems," *Data Mining and Knowledge Discovery*, vol. 24, no. 3, pp. 555–583, 2012.

[343] T. Kulesza, S. Stumpf, M. Burnett, S. Yang, I. Kwan, and W.-K. Wong, "Too much, too little, or just right? Ways explanations impact end users' mental models," in *2013 IEEE Symposium on Visual Languages and Human Centric Computing*. IEEE, 2013, pp. 3–10.

[344] P. Kumaraguru and L. F. Cranor, "Privacy Indexes: A Survey of Westin's Studies," *Institute for Software Research International*, 2005.

[345] J. Tsai, L. F. Cranor, A. Acquisti, and C. M. Fong, "What's it to You? A Survey of Online Privacy Concerns and Risks," *NET Institute Working Paper*, vol. 06, no. 29, pp. 1–20, 2006. DOI: 10.2139/ssrn.941708

[346] A. Westin, L. Harris, and Associates, "Equifax-Harris Consumer Privacy Survey," Equifax Corporate Marketing Department, Atlanta, GA, USA, Tech. Rep., 1994.

[347] A. Westin, L. Harris, and Associates, "Mid-Decade Consumer Privacy Survey," Equifax Corporate Marketing Department, Atlanta, GA, USA, Tech. Rep., 1995.

[348] A. Westin, L. Harris, and Associates, "Consumer Privacy Survey," Equifax Corporate Marketing Department, Atlanta, GA, USA, Tech. Rep., 1996.

[349] K. Renaud and D. Gálvez-Cruz, "Privacy: Aspects, definitions and a multi-faceted privacy preservation approach," in *2010 Information Security for South Africa*, 2010. DOI: 10.1109/ISSA.2010.5588297 pp. 1–8.

[350] A. D. Moore, "Privacy: Its Meaning and Value," *American Philosophical Quarterly*, vol. 40, no. 3, pp. 215–227, 2003.

[351] M. Z. Yao, "Self-Protection of Online Privacy: A Behavioral Approach," in *Privacy Online: Perspectives on Privacy and Self-Disclosure in the Social Web*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, ch. 9, pp. 111–125. DOI: 10.1007/978-3-642-21521-6_9

[352] A. Krishna, "Privacy is a Concern: An Introduction to the Dialogue on Privacy," *Journal of Consumer Psychology*, vol. 30, no. 4, pp. 733–735, 2020.

[353] L. D. Introna, "Privacy and the Computer: Why We Need Privacy in the Information Society," *Metaphilosophy*, vol. 28, no. 3, pp. 259–275, 1997. DOI: 10.1111/1467-9973.00055

[354] P. B. Newell, "Perspectives on privacy," *Journal of Environmental Psychology*, vol. 15, no. 2, pp. 87–104, 1995. DOI: 10.1016/0272-4944(95)90018-7

[355] S. D. Warren and L. D. Brandeis, "The Right to Privacy," *Harvard Law Review*, vol. 4, no. 5, pp. 193–220, 1890.

[356] A. P. Bates, "Privacy — A Useful Concept?" *Social Forces*, vol. 42, no. 4, pp. 429–434, 05 1964. DOI: 10.2307/2574986

[357] J. H. Reiman, "Privacy, Intimacy, and Personhood," *Philosophy & Public Affairs*, vol. 6, no. 1, pp. 26–44, 1976.

[358] R. E. Smith, *Ben Franklin's web site: Privacy and curiosity from Plymouth Rock to the Internet.* Privacy Journal, 2000.

[359] A. L. Allen, *Uneasy Access: Privacy for Women in a Free Society.* Totowa, New Jersey, USA: Rowman & Littlefield, 1988.

[360] P. H. Klopfer and D. I. Rubenstein, "The Concept Privacy and its Biological Basis," *Journal of social Issues*, vol. 33, no. 3, pp. 52–65, 1977.

[361] S. Petronio, *Boundaries of Privacy: Dialectics of Disclosure.* State University of New York Press, 2002.

[362] K. Bräunlich, T. Dienlin, J. Eichenhofer, P. Helm, S. Trepte, R. Grimm, S. Seubert, and C. Gusy, "Linking Loose Ends: An Interdisciplinary Privacy and Communication Model," *New Media & Society*, vol. 23, no. 6, pp. 1443–1464, 2021. DOI: 10.1177/1461444820905045

[363] D. P. Bhave, L. H. Teo, and R. S. Dalal, "Privacy at Work: A Review and a Research Agenda for a Contested Terrain," *Journal of Management*, vol. 46, no. 1, pp. 127–164, 2020. DOI: 10.1177/0149206319878254

[364] E. J. Bloustein, "Privacy as an Aspect of Human Dignity: An Answer to Dean Prosser," *New York University Law Review*, vol. 39, no. 6, pp. 962–1007, 1964.

[365] H. Nissenbaum, "Privacy as Contextual Integrity," *Washington Law Review*, vol. 79, p. 119, 2004.

[366] J. E. Cohen, "Examined Lives: Informational Privacy and the Subject as Object," *Stanford Law Review*, vol. 52, no. 5, pp. 1373–1438, 2000.

[367] D. J. Solove, "Understanding Privacy," in *GWU Law School Public Law Research Paper.* Harvard University Press, 2008, p. 24.

[368] E. T. Hall, *The Hidden Dimension*, 1st ed. New York, NY, USA: Knopf Doubleday Publishing Group, 1966.

[369] A. Senarath and N. A. G. Arachchilage, "Why Developers Cannot Embed Privacy into Software Systems? An Empirical Investigation," in *Proceedings of the 22nd International Conference on Evaluation and Assessment in Software Engineering 2018*, ser. EASE'18. New York, NY, USA: Association for Computing Machinery, 2018. DOI: 10.1145/3210459.3210484 p. 211–216.

[370] M. Peixoto, D. Ferreira, M. Cavalcanti, C. Silva, J. Vilela, J. Araújo, and T. Gorschek, "On Understanding How Developers Perceive and Interpret Privacy Requirements Research Preview," in *Requirements Engineering: Foundation for Software Quality*, N. Madhavji, L. Pasquale, A. Ferrari, and S. Gnesi, Eds. Cham: Springer International Publishing, 2020, pp. 116–123.

[371] A. Acquisti, L. Brandimarte, and G. Loewenstein, "Privacy and human behavior in the age of information," *Science*, vol. 347, no. 6221, pp. 509–514, 2015. DOI: 10.1126/science.aaa1465

[372] S. Jasanoff, "Virtual, visible, and actionable: Data assemblages and the sightlines of justice," *Big Data & Society*, vol. 4, no. 2, pp. 1–15, 2017. DOI: 10.1177/2053951717724477

[373] A. Richardson and A. Rosenfeld, "A survey of interpretability and explainability in human-agent systems," in *Proceedings of the IJCAI/ECAI Workshop on Explainable Artificial Intelligence (XAI 2018)*, 2018, pp. 137–143.

[374] R. M. Groves, F. J. Fowler Jr, M. P. Couper, J. M. Lepkowski, E. Singer, and R. Tourangeau, *Survey methodology*, 2nd ed. Hoboken, NJ, USA: John Wiley & Sons, 2011.

[375] C. Seaman, "Qualitative methods in empirical studies of software engineering," *IEEE Transactions on Software Engineering*, vol. 25, no. 4, pp. 557–572, 1999. DOI: 10.1109/32.799955

[376] J. Cohen, "Weighted Kappa: Nominal Scale Agreement Provision for Scaled Disagreement or Partial Credit." *Psychological bulletin*, vol. 70, no. 4, 1968.

[377] A. Woodruff, V. Pihur, S. Consolvo, L. Brandimarte, and A. Acquisti, "Would a Privacy Fundamentalist Sell Their DNA for $1000...If Nothing Bad Happened as a Result? The Westin Categories, Behavioral Intentions, and Consequences," in *10th Symposium On Usable Privacy and Security (SOUPS 2014)*. Menlo Park, CA: USENIX Association, jul 2014, pp. 1–18.

[378] S. Consolvo, I. E. Smith, T. Matthews, A. LaMarca, J. Tabert, and P. Powledge, "Location Disclosure to Social Relations: Why, When, & What People Want to Share," in *Proceedings*

of the SIGCHI Conference on Human Factors in Computing Systems, ser. CHI '05. New York, NY, USA: Association for Computing Machinery, 2005. DOI: 10.1145/1054972.1054985 p. 81–90.

[379] A. Westin, "Privacy On and Off the Internet: What Consumers Want," in *Privacy and American Business*. New York, NY, USA: Harris Interactive, Feb 2002, pp. 1–126.

[380] L. J. Cronbach, "Coefficient alpha and the internal structure of tests," *Psychometrika*, vol. 16, no. 3, pp. 297–334, Sep 1951. DOI: 10.1007/BF02310555

[381] D. George and P. Mallery, *SPSS for Windows Step by Step: A Simple Study Guide and Reference, 17.0 Update*, 10th ed. USA: Allyn & Bacon, Inc., 2009.

[382] Harris Interactive, "Most People Are „Privacy Pragmatists" Who, While Concerned about Privacy, Will Sometimes Trade It Off for Other Benefits," https://web.archive.org/web/20060208065600/www.harrisinteractive.com/harris_poll/index.asp?PID=365, 2003, Letzter Zugriff: 21.03.2023.

[383] E.-M. Schomakers, C. Lidynia, D. Müllmann, and M. Ziefle, "Internet users' perceptions of information sensitivity – insights from Germany," *International Journal of Information Management*, vol. 46, pp. 142–150, 2019. DOI: 10.1016/j.ijinfomgt.2018.11.018

[384] S. S. Shapiro and M. B. Wilk, "An Analysis of Variance Test for Normality (Complete Samples)," *Biometrika*, vol. 52, no. 3-4, pp. 591–611, 12 1965. DOI: 10.1093/biomet/52.3-4.591

[385] C. Spearman, "The Proof and Measurement of Association between Two Things," *The American Journal of Psychology*, vol. 15, no. 1, pp. 72–101, 1904. DOI: 10.2307/1412159

[386] W. Brunotte, "Data for Research Article *„Privacy Explanations – A Means to End-User Trust"*," 2022. DOI: 10.5281/zenodo.7215560

[387] A. Papenmeier, G. Englebienne, and C. Seifert, "How Model Accuracy and Explanation Fidelity Influence User Trust in AI," in *Proceedings of the IJCAI 2019 Workshop on Explainable Artificial Intelligence (XAI)*, 2019, pp. 94–100.

[388] J. Koskinen, S. Knaapi-Junnila, and M. M. Rantanen, "What if we Had Fair, People-Centred Data Economy Ecosystems?" in *2019 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, 2019. DOI: 10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00100 pp. 329–334.

[389] B. Schafer and L. Edwards, "„I spy, with my little sensor": fair data handling practices for robots between privacy, copyright and security," *Connection Science*, vol. 29, no. 3, pp. 200–209, 2017. DOI: 10.1080/09540091.2017.1318356

[390] W. Brunotte, J. Droste, and K. Schneider, "Context, Content, Consent – How to Design User-Centered Privacy Explanations," in *The 35th International Conference on Software Engineering & Knowledge Engineering*, San Francisco, USA, 2023. DOI: 10.18293/SEKE2023-032. ISSN 2325-9000 pp. 86–89.

[391] J. Jacobsen and L. Meyer, *Praxisbuch Usability und UX: Was jeder wissen sollte, der Websites und Apps entwickelt - bewährte Usability- und UX-Methoden praxisnah erklärt.* Deutschland: Rheinwerk Computing, 2017.

[392] A. Sunyaev, T. Dehling, P. L. Taylor, and K. D. Mandl, "Availability and quality of mobile health app privacy policies," *Journal of the American Medical Informatics Association*, vol. 22, no. e1, pp. e28–e33, 2015.

[393] R. Balebako, F. Schaub, I. Adjerid, A. Acquisti, and L. Cranor, "The impact of timing on the salience of smartphone app privacy notices," in *Proceedings of the 5th Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices*, 2015, pp. 63–74.

[394] F. Kreuter, G.-C. Haas, F. Keusch, S. Bähr, and M. Trappmann, "Collecting Survey and Smartphone Sensor Data With an App: Opportunities and Challenges Around Privacy and Informed Consent," *Social Science Computer Review*, vol. 38, no. 5, pp. 533–549, 2020. DOI: 10.1177/0894439318816389

[395] J. Correia and D. Compeau, "Information Privacy Awareness (IPA): A Review of the Use, Definition and Measurement of IPA," in *50th Hawaii International Conference on System Sciences, HICSS 2017, Hilton Waikoloa Village, Hawaii, USA, January 4-7, 2017*, T. Bui, Ed. ScholarSpace / AIS Electronic Library (AISeL), 2017, pp. 1–10. [Online]. Available: https://hdl.handle.net/10125/41646

[396] J. Lin, S. Amini, J. I. Hong, N. Sadeh, J. Lindqvist, and J. Zhang, "Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing," in *Proceedings of the 2012 ACM conference on ubiquitous computing*, 2012, pp. 501–510.

[397] M. Zanker, "The influence of knowledgeable explanations on users' perception of a recommender system," in *Proceedings of the sixth ACM conference on Recommender systems*, 2012, pp. 269–272.

[398] M. Chromik, M. Eiband, S. T. Völkel, and D. Buschek, "Dark Patterns of Explainability, Transparency, and User Control for Intelligent Systems." in *IUI workshops*, vol. 2327, 2019.

[399] R. Balebako, R. Shay, and L. F. Cranor, "Is your inseam a biometric? evaluating the understandability of mobile privacy notice categories," *CMU, Tech. Rep. CMU-CyLab-13-011*, 2013.

[400] H. Fu and J. Lindqvist, "General Area or Approximate Location? How People Understand Location Permissions," in *Proceedings of the 13th Workshop on Privacy in the Electronic Society*, ser. WPES '14. New York, NY, USA: Association for Computing Machinery, 2014. DOI: 10.1145/2665943.2665957 p. 117–120.

[401] E. J. Johnson, S. Bellman, and G. L. Lohse, "Defaults, Framing and Privacy: Why Opting In-Opting Out," *Marketing Letters*, vol. 13, no. 1, pp. 5–15, Feb 2002. DOI: 10.1023/A:1015044207315

[402] A. Noain-Sánchez, ""Privacy by default" and active "informed consent" by layers," *Journal of Information, Communication and Ethics in Society*, vol. 14, no. 2, pp. 124–138, Jan 2016. DOI: 10.1108/JICES-10-2014-0040

[403] R. R. Hoffman, G. Klein, and S. T. Mueller, "Explaining Explanation For "Explainable AI"," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 62, no. 1, 2018. DOI: 10.1177/1541931218621047 pp. 197–201.

[404] M. T. Ribeiro, S. Singh, and C. Guestrin, "„Why should i trust you?" Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 1135–1144.

[405] K. Martin, "Privacy notices as tabula rasa: An empirical investigation into how complying with a privacy notice is related to meeting privacy expectations online," *Journal of Public Policy & Marketing*, vol. 34, no. 2, pp. 210–227, 2015.

[406] M. Hall, D. Harborne, R. Tomsett, V. Galetic, S. Quintana-Amate, A. Nottle, and A. Preece, "A Systematic Method to Understand Requirements for Explainable AI (XAI) Systems," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 21–27.

[407] H. K. Dam, T. Tran, and A. Ghose, "Explainable Software Analytics," in *Proceedings of the 40th International Conference on Software Engineering: New Ideas and Emerging Results (ICSE-NIER)*. New York, NY, USA: Association for Computing Machinery, 2018. DOI: 10.1145/3183399.3183424 pp. 53–56.

[408] A. D. Preece, D. Harborne, D. Braines, R. Tomsett, and S. Chakraborty, "Stakeholders in Explainable AI," *CoRR*, vol. abs/1810.00184, 2018. [Online]. Available: http://arxiv.org/abs/1810.00184

[409] N. Tintarev and J. Masthoff, "Effective explanations of recommendations: user-centered design," in *Proceedings of the 2007 ACM conference on Recommender systems*, 2007, pp. 153–156.

[410] A. Weller, "Transparency: Motivations and Challenges," in *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*. Springer International Publishing, 2019, pp. 23–40. DOI: 10.1007/978-3-030-28954-6_2

[411] M. E. Kaminski and G. Malgieri, "Multi-Layered Explanations from Algorithmic Impact Assessments in the GDPR," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, ser. FAT* '20. New York, NY, USA: Association for Computing Machinery, 2020. DOI: 10.1145/3351095.3372875 p. 68–79.

[412] E. Weber, J. Van Bouwel, and R. Vanderbeeken, "Forms of causal explanation," *Foundations of Science*, vol. 10, no. 4, pp. 437–454, 2005.

[413] J. Y. Halpern and J. Pearl, "Causes and Explanations: A Structural-Model Approach. Part II: Explanations," *The British Journal for the Philosophy of Science*, vol. 56, no. 4, pp. 889–911, 2005. DOI: 10.1093/bjps/axi148

[414] R. M. Byrne, "Counterfactuals in Explainable Artificial Intelligence (XAI): Evidence from Human Reasoning," in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19)*, 2019. DOI: 10.24963/ijcai.2019/876 pp. 6276–6282.

[415] W. Bechtel, "Levels of description and explanation in cognitive science," *Minds and Machines*, vol. 4, no. 1, pp. 1–25, 1994.

[416] W. Bechtel and A. Abrahamsen, "Explanation: A Mechanist Alternative," *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, vol. 36, no. 2, pp. 421–441, 2005. DOI: 10.1016/j.shpsc.2005.03.010 Mechanisms in biology.

[417] B. Y. Lim and A. K. Dey, "Assessing demand for intelligibility in context-aware applications," in *Proceedings of the 11th international conference on Ubiquitous computing*, 2009, pp. 195–204.

[418] P. Pu and L. Chen, "Trust-inspiring explanation interfaces for recommender systems," *Knowledge-Based Systems*, vol. 20, no. 6, pp. 542–556, 2007.

[419] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi, "A survey of methods for explaining black box models," *ACM computing surveys (CSUR)*, vol. 51, no. 5, pp. 1–42, 2018.

[420] S. Wachter, B. Mittelstadt, and C. Russell, "Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR ," *Harvard Journal of Law & Technology (Harvard JOLT)*, vol. 31, no. 2, pp. 841–888, 2018.

[421] V. Rajlich, J. Doran, and R. Gudla, "Layered Explanations of Software: A Methodology for Program Comprehension," in *Proceedings 1994 IEEE 3rd Workshop on Program Comprehension- WPC '94*, 1994. DOI: 10.1109/WPC.1994.341248 pp. 46–52.

[422] F. Wilcoxon, "Individual Comparisons by Ranking Methods," *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945.

[423] H. B. Mann and D. R. Whitney, "On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other," *The Annals of Mathematical Statistics*, vol. 18, no. 1, pp. 50 – 60, 1947. DOI: 10.1214/aoms/1177730491

[424] O. Ulvi, A. Karamehic-Muratovic, M. Baghbanzadeh, A. Bashir, J. Smith, and U. Haque, "Social Media Use and Mental Health: A Global Analysis," *Epidemiologia*, vol. 3, no. 1, pp. 11–25, 2022. DOI: 10.3390/epidemiologia3010002

[425] I. Anger and C. Kittl, "Measuring Influence on Twitter," in *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, ser. i-KNOW '11. New York, NY, USA: Association for Computing Machinery, 2011. DOI: 10.1145/2024288.2024326

[426] H. Khan, U. Hengartner, and D. Vogel, "Usability and Security Perceptions of Implicit Authentication: Convenient, Secure, Sometimes Annoying," in *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*. Ottawa: USENIX Association, 2015, pp. 225–239.

[427] M. Harbach, E. von Zezschwitz, A. Fichtner, A. D. Luca, and M. Smith, "Its a Hard Lock Life: A Field Study of Smartphone (Un)Locking Behavior and Risk Perception," in *10th Symposium On Usable Privacy and Security (SOUPS 2014)*. Menlo Park, CA: USENIX Association, 2014, pp. 213–230.

[428] D. Klitou, *Privacy-Invading Technologies and Privacy by Design*, 1st ed. The Hague, NL: T.M.C. Asser Press, The Hague, 2014.

[429] S. Ahern, D. Eckles, N. S. Good, S. King, M. Naaman, and R. Nair, "Over-Exposed? Privacy Patterns and Considerations in Online and Mobile Photo Sharing," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '07. New York, NY, USA: Association for Computing Machinery, 2007. DOI: 10.1145/1240624.1240683 p. 357–366.

[430] D. E. Pozen, "Privacy-Privacy Tradeoffs," *The University of Chicago Law Review*, vol. 83, no. 1, pp. 221–247, 2016.

[431] D. E. Bambauer, "Privacy versus Security," *Journal of Criminal Law and Criminology*, vol. 103, no. 3, pp. 667–684, 2013.

[432] C. J. Bennet, *The Privacy Advocates: Resisting The Spread Of Surveillance*. Cambridge, Massachusetts, USA: The MIT Press, 2008.

[433] T. Ogura, "Electronic government and surveillance-oriented society," in *Theorizing surveillance: The Panopticon and Beyond*. Cullompton, UK: Willan Publishing, 2006, pp. 284–309.

[434] D. Lyon, "Surveillance technology and surveillance society," *Modernity and technology*, pp. 161–184, 2003.

[435] D. Lyon, *Surveillance society: Monitoring everyday life*. Buckingham, Philadelphia, USA: Open University Press, 2001.

[436] R. Tate, "Google CEO: Secrets Are for Filthy People," https://www.gawker.com/5419271/google-ceo-secrets-are-for-filthy-people, Letzter Zugriff: 31.10.2021.

[437] H. Kubicek, "Vertrauen durch Sicherheit — Vertrauen in Sicherheit. Annäherung an ein schwieriges Verhältnis," in *Informationelles Vertrauen für die Informationsgesellschaft*, D. Klumpp, H. Kubicek, A. Roßnagel, and W. Schulz, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 17–35. DOI: 10.1007/978-3-540-77670-3_2

[438] D. H. McKnight, "Trust in Information Technology," in *The Blackwell Encyclopedia of Management. Vol. 7 Management Information Systems*. Malden, MA, USA: Blackwell Publishing, 2005, pp. 329–331.

[439] F. Schaub, R. Balebako, and L. F. Cranor, "Designing Effective Privacy Notices and Controls," *IEEE Internet Computing*, vol. 21, no. 3, pp. 70–77, 2017. DOI: 10.1109/MIC.2017.75

[440] C. Bowman, A. Gesher, J. K. Grant, and D. Slate, *The Architecture of Privacy*, 1st ed. Sebastopol, CA, USA: O'Reilly Media, Inc., 2015.

[441] E. C. Groen, D. Feth, S. Polst, J. Tolsdorf, S. Wiefling, L. L. Iacono, and H. Schmitt, "Achieving Usable Security and Privacy Through Human-Centered Design," in *Human Factors in Privacy Research*, N. Gerber, A. Stöver, and K. Marky, Eds. Cham: Springer International Publishing, 2023, pp. 83–113. DOI: 10.1007/978-3-031-28643-8_5

[442] P. Ohm, "Foreword," in *The Architecture of Privacy*, 1st ed. Sebastopol, CA, USA: O'Reilly Media, Inc., 2015, pp. 9–11.

[443] A. Seffah and E. Metzker, "The Obstacles and Myths of Usability and Software Engineering," *Commun. ACM*, vol. 47, no. 12, p. 71–76, dec 2004. DOI: 10.1145/1035134.1035136

[444] W. Brunotte, "Security Code Clone Detection entwickelt als Eclipse Plugin," Master's thesis, Leibniz Universität Hannover, Fachgebiet Software Engineering, Hannover, Germany, June 2018.

[445] R. Jochem, *Was kostet Qualität? – Wirtschaftlichkeit von Qualität ermitteln.* München, Deutschland: Carl Hanser Verlag GmbH & Co. KG, 2010.

[446] A. N. Venturelli, "A cautionary contribution to the philosophy of explanation in the cognitive neurosciences," *Minds and Machines*, vol. 26, no. 3, pp. 259–285, 2016.

[447] J. Sliwinski, M. Strobel, and Y. Zick, "A characterization of monotone influence measures for data classification," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2017)*, 2017, pp. 48–52.

[448] C. Brinton, "A Framework for Explanation of Machine Learning Decisions," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2017)*, 2017, pp. 14–18.

[449] R. Pierrard, J.-P. Poli, and C. Hudelot, "A New Approach for Explainable Multiple Organ Annotation with Few Data," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 101–107.

[450] P. B. De Laat, "Algorithmic decision-making based on machine learning from Big Data: Can transparency restore accountability?" *Philosophy & Technology*, vol. 31, no. 4, pp. 525–541, 2018. DOI: 10.1007/s13347-017-0293-z

[451] I. Monteath and R. Sheh, "Assisted and incremental medical diagnosis using explainable artificial intelligence," in *Proceedings of the IJCAI/ECAI Workshop on Explainable Artificial Intelligence (XAI 2018)*, 2018, pp. 104–108.

[452] U. Ehsan, P. Tambwekar, L. Chan, B. Harrison, and M. O. Riedl, "Automated rationale generation: a technique for explainable AI and its effects on human perceptions," in *Proceedings of the 24th International Conference on Intelligent User Interfaces*, 2019, pp. 263–274.

[453] A. Aggarwal, P. Lohia, S. Nagar, K. Dey, and D. Saha, "Black Box Fairness Testing of Machine Learning Models," in *Proceedings of the 2019 27th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*. New York, NY, USA: Association for Computing Machinery, 2019. DOI: 10.1145/3338906.3338937 p. 625–635.

[454] K. L. Theurer, "Compositional explanatory relations and mechanistic reduction," *Minds and Machines*, vol. 23, no. 3, pp. 287–307, 2013.

[455] D. J. Buller, "Confirmation and the computational paradigm (or: Why do you think they call it artificial intelligence?)," *Minds and Machines*, vol. 3, no. 2, pp. 155–181, 1993.

[456] A. Lucic, H. Haned, and d. R. Maarten, "Contrastive Explanations for Large Errors in Retail Forecasting Predictions through Monte Carlo Simulations," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 66–72.

[457] K. Sokol and P. A. Flach, "Conversational Explanations of Machine Learning Predictions Through Class-contrastive Counterfactual Statements." in *Proceedings of the IJCAI/ECAI Workshop on Explainable Artificial Intelligence (XAI 2018)*, 2018, pp. 5785–5786.

[458] M. L. Olson, L. Neal, F. Li, and W.-K. Wong, "Counterfactual States for Atari Agents via Generative Deep Learning," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 87–93.

[459] L. Hiley, A. Preece, Y. Hicks, D. Marshall, and H. Taylor, "Discriminating spatial and temporal relevance in deep Taylor decompositions for explainable activity recognition," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 35–39.

[460] Y. Coppens, K. Efthymiadis, T. Lenaerts, A. Nowé, T. Miller, R. Weber, and D. Magazzeni, "Distilling deep reinforcement learning policies in soft decision trees," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 1–6.

[461] S. T. Mckinlay, "Evidence, explanation and predictive data modelling," *Philosophy & Technology*, vol. 30, no. 4, pp. 461–473, 2017.

[462] T. Miller, P. Howe, and L. Sonenberg, "Explainable AI: Beware of Inmates Running the Asylum. Or: How I Learnt to Stop Worrying and Love the Social and Behavioural Sciences," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2017)*, 2017, pp. 36–42.

[463] P. S. Kumar, M. Saravanan, and S. Suresh, "Explainable Classification Using Clustering in Deep Learning Models," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 115–121.

[464] M. Fox, D. Long, and D. Magazzeni, "Explainable Planning," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2017)*, 2017, pp. 24–30.

[465] P. Madumal, T. Miller, L. Sonenberg, and F. Vetere, "Explainable Reinforcement Learning Through a Causal Lens," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 73–79.

[466] F. Gutiérrez, S. Charleer, R. De Croon, N. N. Htun, G. Goetschalckx, and K. Verbert, "Explaining and exploring job recommendations: a user-driven approach for interacting with knowledge-based job recommender systems," in *Proceedings of the 13th ACM Conference on Recommender Systems*, 2019, pp. 60–68.

[467] R. O. Weber, H. Hong, and P. Goel, "Explaining Citation Recommendations: Abstracts or Full Texts?" in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 136–142.

[468] N. Angius and G. Tamburrini, "Explaining engineered computing systems' behaviour: the role of abstraction and idealization," *Philosophy & Technology*, vol. 30, no. 2, pp. 239–258, 2017.

[469] J. Dodge, Q. V. Liao, Y. Zhang, R. K. Bellamy, and C. Dugan, "Explaining models: an empirical study of how explanations impact fairness judgment," in *Proceedings of the 24th International Conference on Intelligent User Interfaces*, 2019, pp. 275–285.

[470] L. Chen and F. Wang, "Explaining recommendations based on feature sentiments in product reviews," in *Proceedings of the 22nd international conference on intelligent user interfaces*, 2017, pp. 17–28.

[471] 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*, 2019, pp. 391–396.

[472] O. Biran and C. Cotton, "Explanation and Justification in Machine Learning: A Survey," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2017)*, 2017, pp. 8–13.

[473] A. Gopnik, "Explanation as orgasm," *Minds and machines*, vol. 8, no. 1, pp. 101–118, 1998.

[474] J. Fernández, "Explanation by computer simulation in cognitive science," *Minds and Machines*, vol. 13, no. 2, pp. 269–284, 2003.

[475] H. Johnson and P. Johnson, "Explanation facilities and interactive systems," in *Proceedings of the 1st International Conference on Intelligent User Interfaces (IUI)*. New York, NY, USA: Association for Computing Machinery, 1993. DOI: 10.1145/169891.169951 pp. 159–166.

[476] W. F. Brewer, C. A. Chinn, and A. Samarapungavan, "Explanation in scientists and children," *Minds and Machines*, vol. 8, no. 1, pp. 119–136, 1998.

[477] A. Rueger, "Explanations at Multiple Levels," *Minds and Machines*, vol. 11, no. 4, pp. 503–520, 2001.

[478] J. De Winter, "Explanations in software engineering: The pragmatic point of view," *Minds and Machines*, vol. 20, no. 2, pp. 277–289, 2010. DOI: 10.1007/s11023-010-9190-2

[479] N. Tintarev, "Explanations of Recommendations," in *Proceedings of the 2007 ACM Conference on Recommender Systems*. New York, NY, USA: Association for Computing Machinery, 2007. DOI: 10.1145/1297231.1297275 pp. 203–206.

[480] C. Lucero, B. Coronado, O. Hui, and D. S. Lange, "Exploring explainable artificial intelligence and autonomy through provenance," in *Proceedings of the IJCAI/ECAI Workshop on Explainable Artificial Intelligence (XAI 2018)*, 2018, p. 85.

[481] V. Putnam and C. Conati, "Exploring the Need for Explainable Artificial Intelligence (XAI) in Intelligent Tutoring Systems (ITS)." in *IUI Workshops*, vol. 19, 2019.

[482] A. J. Ko and B. A. Myers, "Extracting and answering why and why not questions about Java program output," *ACM Transactions on Software Engineering and Methodology (TOSEM)*, vol. 20, no. 2, pp. 1–36, 2010.

[483] M. Veale and R. Binns, "Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data," *Big Data & Society*, vol. 4, no. 2, pp. 1–17, 2017. DOI: 10.1177/2053951717743530

[484] M. Hosseini, A. Shahri, K. Phalp, and R. Ali, "Four reference models for transparency requirements in information systems," *Requirements Engineering*, vol. 23, no. 2, pp. 251–275, 2018.

[485] G. J. Nalepa, M. van Otterlo, S. Bobek, and M. Atzmueller, "From context mediation to declarative values and explainability," in *Proceedings of the IJCAI/ECAI Workshop on Explainable Artificial Intelligence (XAI 2018)*, 2018.

[486] S. J. Green, P. Lamere, J. Alexander, F. Maillet, S. Kirk, J. Holt, J. Bourque, and X.-W. Mak, "Generating transparent, steerable recommendations from textual descriptions of items," in *Proceedings of the third ACM conference on Recommender systems*, 2009, pp. 281–284.

[487] J. Schaffer, P. Giridhar, D. Jones, T. Höllerer, T. Abdelzaher, and J. O'donovan, "Getting the message? A study of explanation interfaces for microblog data analysis," in *Proceedings of the 20th international conference on intelligent user interfaces*, 2015, pp. 345–356.

[488] J. C.-T. Ho, "How biased is the sample? Reverse engineering the ranking algorithm of Facebook's Graph application programming interface," *Big Data & Society*, vol. 7, no. 1, pp. 1–15, 2020. DOI: 10.1177/2053951720905874

[489] T. Barik, D. Ford, E. Murphy-Hill, and C. Parnin, "How should compilers explain problems to developers?" in *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, 2018, pp. 633–643.

[490] J. Burrell, "How the machine 'thinks': Understanding opacity in machine learning algorithms," *Big Data & Society*, vol. 3, no. 1, pp. 1–12, 2016. DOI: 10.1177/2053951715622512

[491] J. Schaffer, J. O'Donovan, J. Michaelis, A. Raglin, and T. Höllerer, "I can do better than your AI: expertise and explanations," in *Proceedings of the 24th International Conference on Intelligent User Interfaces*, 2019, pp. 240–251.

[492] H. Liu, J. Wen, L. Jing, J. Yu, X. Zhang, and M. Zhang, "In2Rec: Influence-based interpretable recommendation," in *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 2019, pp. 1803–1812.

[493] B. Ghosh, D. Malioutov, and K. S. Meel, "Interpretable Classification Rules in Relaxed Logical Form," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 14–20.

[494] T. Huber, D. Schiller, and E. André, "Introducing Selective Layer Wise Relevance Propagation to Dueling Deep Q learning," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019))*, 2019.

[495] X. Watts and F. Lécué, "Local Score Dependent Model Explanation for Time Dependent Covariates," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 129–135.

[496] L. Michael, "Machine Coaching," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 80–86.

[497] R. McClamrock, "Marr's three levels: A re-evaluation," *Minds and Machines*, vol. 1, no. 2, pp. 185–196, 1991.

[498] M. T. Stuart and N. J. Nersessian, "Peeking inside the black box: a new kind of scientific visualization," *Minds and Machines*, vol. 29, no. 1, pp. 87–107, 2019. DOI: 10.1007/s11023-018-9484-3

[499] T. Kulesza, M. Burnett, W.-K. Wong, and S. Stumpf, "Principles of explanatory debugging to personalize interactive machine learning," in *Proceedings of the 20th international conference on intelligent user interfaces*, 2015, pp. 126–137.

[500] P. Sawyer, N. Bencomo, J. Whittle, E. Letier, and A. Finkelstein, "Requirements-aware systems: A research agenda for re for self-adaptive systems," in *2010 18th IEEE International Requirements Engineering Conference*. IEEE, 2010, pp. 95–103.

[501] A. Vogelsang and M. Borg, "Requirements engineering for machine learning: Perspectives from data scientists," in *2019 IEEE 27th International Requirements Engineering Conference Workshops (REW)*. IEEE, 2019, pp. 245–251.

[502] C. Zednik, "Solving the black box problem: A normative framework for explainable artificial intelligence," *Philosophy & Technology*, pp. 1–24, 2019. DOI: 10.1007/s13347-019-00382-7

[503] R. Häuslschmid, M. von Buelow, B. Pfleging, and A. Butz, "Supporting Trust in autonomous driving," in *Proceedings of the 22nd international conference on intelligent user interfaces*, 2017, pp. 319–329.

[504] J. Vig, S. Sen, and J. Riedl, "Tagsplanations: Explaining Recommendations Using Tags," in *Proceedings of the 14th International Conference on Intelligent User Interfaces (IUI.* New York, NY, USA: Association for Computing Machinery, 2009. DOI: 10.1145/1502650.1502661 pp. 47–56.

[505] K. Ehrlich, S. E. Kirk, J. Patterson, J. C. Rasmussen, S. I. Ross, and D. M. Gruen, "Taking Advice from Intelligent Systems: The Double-Edged Sword of Explanations," in *Proceedings of the 16th International Conference on Intelligent User Interfaces (IUI)*. New York, NY, USA: Association for Computing Machinery, 2011. DOI: 10.1145/1943403.1943424 pp. 125–134.

[506] V. Dominguez, P. Messina, I. Donoso-Guzmán, and D. Parra, "The effect of explanations and algorithmic accuracy on visual recommender systems of artistic images," in *Proceedings of the 24th International Conference on Intelligent User Interfaces*, 2019, pp. 408–416.

[507] B. D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, "The ethics of algorithms: Mapping the debate," *Big Data & Society*, vol. 3, no. 2, pp. 1–21, 2016. DOI: 10.1177/2053951716679679

[508] L. Floridi, "The method of levels of abstraction," *Minds and machines*, vol. 18, no. 3, pp. 303–329, 2008.

[509] A. Páez, "The Pragmatic Turn in Explainable Artificial Intelligence (XAI)," *Minds & Machines*, vol. 29, pp. 441–459, 2019. DOI: 10.1007/s11023-019-09502-w

[510] R. A. Wilson and F. Keil, "The shadows and shallows of explanation," *Minds and machines*, vol. 8, no. 1, pp. 137–159, 1998.

[511] M. T. Keane and E. M. Kenny, "The twin-system approach as one generic solution for XAI: An overview of ANN-CBR twins for explaining deep learning," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 54–58.

[512] J. Clos, N. Wiratunga, and S. Massie, "Towards Explainable Text Classification by Jointly Learning Lexicon and Modifier Terms," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2017)*, 2017, pp. 19–23.

[513] H. Wicaksono, C. Sammut, and R. Sheh, "Towards Explainable Tool Creation by a Robot," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2017)*, 2017, pp. 63–67.

[514] V. Putnam, L. Riegel, and C. Conati, "Towards Personalized XAI: A Case Study in Intelligent Tutoring Systems," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 108–114.

[515] R. Borgo, M. Cashmore, and D. Magazzeni, "Towards providing explanations for AI planner decisions," in *Proceedings of the IJCAI/ECAI Workshop on Explainable Artificial Intelligence (XAI 2018)*, 2018.

[516] J. Zhou and F. Chen, "Towards Trustworthy Human-AI Teaming under Uncertainty," in *Proceedings of the IJCAI Workshop on Explainable Artificial Intelligence (XAI 2019)*, 2019, pp. 143–147.

[517] J. Zerilli, A. Knott, J. Maclaurin, and C. Gavaghan, "Transparency in algorithmic and human decision-making: is there a double standard?" *Philosophy & Technology*, vol. 32, no. 4, pp. 661–683, 2019. DOI: 10.1007/s13347-018-0330-6

[518] H. Felzmann, E. F. Villaronga, C. Lutz, and A. Tamò-Larrieux, "Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns," *Big Data & Society*, vol. 6, no. 1, pp. 1–14, 2019. DOI: 10.1177/2053951719860542

[519] X. He, T. Chen, M.-Y. Kan, and X. Chen, "Trirank: Review-aware explainable recommendation by modeling aspects," in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, 2015, pp. 1661–1670.

[520] P. Pu and L. Chen, "Trust building with explanation interfaces," in *Proceedings of the 11th international conference on Intelligent user interfaces*, 2006, pp. 93–100.

[521] J. Drozdal, J. Weisz, D. Wang, G. Dass, B. Yao, C. Zhao, M. Muller, L. Ju, and H. Su, "Trust in automl: Exploring information needs for establishing trust in automated machine learning systems," in *Proceedings of the 25th International Conference on Intelligent User Interfaces*, 2020, pp. 297–307.

[522] D. Holliday, S. Wilson, and S. Stumpf, "User trust in intelligent systems: A journey over time," in *Proceedings of the 21st International Conference on Intelligent User Interfaces*, 2016, pp. 164–168.

[523] N. F. Rajani and R. J. Mooney, "Using Explanations to Improve Ensembling of Visual Question Answering Systems," in *Proceedings of the IJCAI 2017 Workshop on Explainable Artificial Intelligence (XAI)*, 2017, pp. 43–47.

[524] D. van Eck, "Validating function-based design methods: An explanationist perspective," *Philosophy & Technology*, vol. 28, no. 4, pp. 511–531, 2015.

[525] P. Madumal, T. Miller, L. Sonenberg, and F. Vetere, "A Grounded Interaction Protocol for Explainable Artificial Intelligence," in *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. Richland, SC, USA: International Foundation for Autonomous Agents and Multiagent Systems, 2019. DOI: 10.5555/3306127.3331801 pp. 1033–1041.

[526] D. A. Wilkenfeld, "Functional explaining: A new approach to the philosophy of explanation," *Synthese*, vol. 191, pp. 3367–3391, 2014. DOI: 10.1007/s11229-014-0452-z

[527] L. Viganò and D. Magazzeni, "Explainable Security," in *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS PW)*, 2020. DOI: 10.1109/EuroSPW51379.2020.00045 pp. 293–300.

[528] M. Milkowski, "A Mechanistic Account of Computational Explanation in Cognitive Science," in *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, M. Knauff, M. Pauen, N. Sebanz, and I. Wachsmuth, Eds. Cognitive Science Society, 2013, pp. 3050–3055. [Online]. Available: https://mindmodeling.org/cogsci2013/papers/0545/index.html

[529] D. Billsus and M. J. Pazzani, "A personal news agent that talks, learns and explains," in *Proceedings of the third annual conference on Autonomous Agents*, 1999, pp. 268–275.

[530] M. Harbers, K. van den Bosch, and J.-J. C. Meyer, "A study into preferred explanations of virtual agent behavior," in *International Workshop on Intelligent Virtual Agents*. Springer, 2009, pp. 132–145.

[531] M.-A. Clinciu and H. Hastie, "A Survey of Explainable AI Terminology," in *Proceedings of the 1st Workshop on Interactive Natural Language Technology for Explainable Artificial Intelligence (NL4XAI 2019)*. Association for Computational Linguistics, 2019. DOI: 10.18653/v1/W19-8403 pp. 8–13.

[532] A. Datta, S. Sen, and Y. Zick, "Algorithmic transparency via quantitative input influence: Theory and experiments with learning systems," in *2016 IEEE symposium on security and privacy (SP)*. IEEE, 2016, pp. 598–617.

[533] I. Lage, E. Chen, J. He, M. Narayanan, B. Kim, S. Gershman, and F. Doshi-Velez, "An Evaluation of the Human-Interpretability of Explanation," *CoRR*, vol. abs/1902.00006, 2019. DOI: 10.48550/arXiv.1902.00006

[534] L. M. Cysneiros and V. M. B. Werneck, "An Initial Analysis on How Software Transparency and Trust Influence each other." in *WER*. Citeseer, 2009.

[535] K. Čyras, D. Letsios, R. Misener, and F. Toni, "Argumentation for Explainable Scheduling," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, pp. 2752–2759, Jul. 2019. DOI: 10.1609/aaai.v33i01.33012752

[536] K. Darlington, "Aspects of intelligent systems explanation," *Universal Journal of Control and Automation*, vol. 1, no. 2, pp. 40–51, 2013. DOI: 10.13189/ujca.2013.010204

[537] Y. Fukuchi, M. Osawa, H. Yamakawa, and M. Imai, "Autonomous self-explanation of behavior for interactive reinforcement learning agents," in *Proceedings of the 5th International Conference on Human Agent Interaction*, 2017, pp. 97–101.

[538] Z. Zeng, C. Miao, C. Leung, and J. J. Chin, "Building more explainable artificial intelligence with argumentation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, 2018.

[539] T. Laugel, M.-J. Lesot, C. Marsala, X. Renard, and M. Detyniecki, "Comparison-based inverse classification for interpretability in machine learning," in *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*. Springer, 2018, pp. 100–111.

[540] A. A. Freitas, "Comprehensible Classification Models: A Position Paper," *SIGKDD Explorations Newsletter*, vol. 15, no. 1, pp. 1–10, 2014. DOI: 10.1145/2594473.2594475

[541] M. Sato, K. Nagatani, T. Sonoda, Q. Zhang, and T. Ohkuma, "Context Style Explanation for Recommender Systems," *Journal of Information Processing*, vol. 27, pp. 720–729, 2019. DOI: 10.2197/ipsjjip.27.720

[542] M. Hosseini, A. Shahri, K. Phalp, and R. Ali, "Crowdsourcing transparency requirements through structured feedback and social adaptation," in *2016 IEEE Tenth International Conference on Research Challenges in Information Science (RCIS)*. IEEE, 2016, pp. 1–6.

[543] A. J. Ko and B. A. Myers, "Debugging Reinvented: Asking and Answering Why and Why Not Questions about Program Behavior," in *Proceedings of the 30th International Conference on Software Engineering*, ser. ICSE '08. New York, NY, USA: Association for Computing Machinery, 2008. DOI: 10.1145/1368088.1368130 p. 301–310.

[544] R. Sheh and I. Monteath, "Defining Explainable AI for Requirements Analysis," *KI-Künstliche Intelligenz*, vol. 32, no. 4, pp. 261–266, 2018. DOI: 10.1007/s13218-018-0559-3

[545] U. Chajewska and J. Y. Halpern, "Defining Explanation in Probabilistic Systems," in *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1997, p. 62–71.

[546] N. Tintarev and J. Masthoff, "Designing and evaluating explanations for recommender systems," in *Recommender systems handbook*. Springer, 2011, pp. 479–510.

[547] T. Hirsch, K. Merced, S. Narayanan, Z. E. Imel, and D. C. Atkins, "Designing Contestability: Interaction Design, Machine Learning, and Mental Health," in *Proceedings of the 2017*

Conference on Designing Interactive Systems, ser. DIS '17. New York, NY, USA: Association for Computing Machinery, 2017. DOI: 10.1145/3064663.3064703 p. 95–99.

[548] D. Wang, Q. Yang, A. Abdul, and B. Y. Lim, "Designing Theory-Driven User-Centric Explainable AI," in *Proceedings of the 2019 Conference on Human Factors in Computing Systems (CHI)*. New York, NY, USA: Association for Computing Machinery, 2019. DOI: 10.1145/3290605.3300831 pp. 1–15.

[549] D. Walton, "Dialogical Models of Explanation." *Association for the Advancement of Artificial Intelligence*, vol. 2007, pp. 1–9, 2007.

[550] R. Sheh, "Different XAI for different HRI," in *AAAI Fall Symposium*. AAAI Press, 2017, pp. 114–117.

[551] M. Ter Hoeve, M. Heruer, D. Odijk, A. Schuth, and M. de Rijke, "Do news consumers want explanations for personalized news rankings," in *FATREC Workshop on Responsible Recommendation Proceedings*, 2017.

[552] D. M. Truxillo, T. E. Bodner, M. Bertolino, T. N. Bauer, and C. A. Yonce, "Effects of explanations on applicant reactions: A meta-analytic review," *International Journal of Selection and Assessment*, vol. 17, no. 4, pp. 346–361, 2009.

[553] B. Goodman and S. Flaxman, "European Union Regulations on Algorithmic Decision-Making and a „Right to Explanation"," *AI Magazine*, vol. 38, no. 3, pp. 50–57, 2017. DOI: 10.1609/aimag.v38i3.2741

[554] N. Tintarev and J. Masthoff, "Evaluating the effectiveness of explanations for recommender systems," *User Modeling and User-Adapted Interaction*, vol. 22, no. 4-5, pp. 399–439, 2012.

[555] S. Anjomshoae, A. Najjar, D. Calvaresi, and K. Främling, "Explainable Agents and Robots: Results from a Systematic Literature Review," in *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. Richland, SC, USA: International Foundation for Autonomous Agents and Multiagent Systems, 2019. DOI: 10.5555/3306127.3331806 p. 1078–1088.

[556] J. Zhu, A. Liapis, S. Risi, R. Bidarra, and G. M. Youngblood, "Explainable AI for designers: A human-centered perspective on mixed-initiative co-creation," in *IEEE Conference on Computational Intelligence and Games (CIG)*. IEEE, 2018. DOI: 10.1109/CIG.2018.8490433 pp. 1–8.

[557] F. Kamiran and I. Žliobaitė, "Explainable and non-explainable discrimination in classification," in *Discrimination and Privacy in the Information Society*. Springer, 2013, pp. 155–170.

[558] S. M. Mathews, "Explainable Artificial Intelligence Applications in NLP, Biomedical, and Malware Classification: A Literature Review," in *Intelligent Computing – Proceedings of the Computing Conference*. Springer, 2019. DOI: 10.1007/978-3-030-22868-2_90 pp. 1269–1292.

[559] G. Vilone and L. Longo, "Notions of Explainability and Evaluation Approaches for Explainable Artificial Intelligence," *Information Fusion*, vol. 76, pp. 89–106, 2021. DOI: 10.1016/j.inffus.2021.05.009

[560] N. Wang, H. Wang, Y. Jia, and Y. Yin, "Explainable recommendation via multi-task learning in opinionated text data," in *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 2018, pp. 165–174.

[561] J. L. Herlocker, J. A. Konstan, and J. Riedl, "Explaining Collaborative Filtering Recommendations," in *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work (CSCW)*. New York, NY, USA: Association for Computing Machinery, 2000. DOI: 10.1145/358916.358995 pp. 241–250.

[562] H.-F. Cheng, R. Wang, Z. Zhang, F. O'Connell, T. Gray, F. M. Harper, and H. Zhu, "Explaining decision-making algorithms through UI: Strategies to help non-expert stakeholders," in *Proceedings of the 2019 CHI conference on human factors in computing systems*. New York, NY, USA: Association for Computing Machinery, 2019. DOI: 10.1145/3290605.3300789 pp. 1–12.

[563] L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, and L. Kagal, "Explaining Explanations: An Overview of Interpretability of Machine Learning," in *IEEE 5th International Conference on Data Science and Advanced Analytics DSAA*, 2018. DOI: 10.1109/DSAA.2018.00018 pp. 80–89.

[564] B. D. Mittelstadt, C. Russell, and S. Wachter, "Explaining Explanations in AI," in *Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency*. New York, NY, USA: Association for Computing Machinery, 2019. DOI: 10.1145/3287560.3287574 pp. 279–288.

[565] L. H. Gilpin, C. Testart, N. Fruchter, and J. Adebayo, "Explaining Explanations to Society," in *NIPS Workshop on Ethical, Social and Governance Issues in AI*, 2018, pp. 1–6.

[566] M. Bilgic and R. J. Mooney, "Explaining recommendations: Satisfaction vs. promotion," in *Beyond Personalization Workshop, IUI*, vol. 5, 2005, p. 153.

[567] K. Cotter, J. Cho, and E. Rader, "Explaining the news feed algorithm: An analysis of the News Feed FYI"blog," in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA)*. New York, NY, USA: Association for Computing Machinery, 2017. DOI: 10.1145/3027063.3053114 pp. 1553–1560.

[568] T. Lombrozo and N. Z. Gwynne, "Explanation and inference: Mechanistic and functional explanations guide property generalization," *Frontiers in Human Neuroscience*, vol. 8, p. 700, 2014.

[569] K. McCain, "Explanation and the nature of scientific knowledge," *Science & Education*, vol. 24, no. 7, pp. 827–854, 2015.

[570] F. C. Keil, "Explanation and Understanding," *Annual Review of Psychology*, vol. 57, no. 1, pp. 227–254, 2006. DOI: 10.1146/annurev.psych.57.102904.190100

[571] F. Sørmo, J. Cassens, and A. Aamodt, "Explanation in case-based reasoning–perspectives and goals," *Artificial Intelligence Review*, vol. 24, no. 2, pp. 109–143, 2005. DOI: 10.1007/s10462-005-4607-7

[572] G. Ras, v. G. Marcel, and P. Haselager, "Explanation methods in deep learning: Users, values, concerns and challenges," in *Explainable and Interpretable Models in Computer Vision and Machine Learning*. Springer, 2018, pp. 19–36. DOI: 10.1007/978-3-319-98131-4_2

[573] E. I. Sklar and M. Q. Azhar, "Explanation Through Argumentation," in *Proceedings of the 6th International Conference on Human-Agent Interaction (HAI)*. New York, NY, USA: Association for Computing Machinery, 2018. DOI: 10.1145/3284432.3284470 pp. 277–285.

[574] D. Walton, "Explanations and Arguments Based on Practical Reasoning." in *ExaCt*, 2009, pp. 72–83.

[575] E. Rader, K. Cotter, and J. Cho, "Explanations as Mechanisms for Supporting Algorithmic Transparency," in *Proceedings of the 2018 Conference on Human Factors in Computing Systems (CHI)*. New York, NY, USA: Association for Computing Machinery, 2018. DOI: 10.1145/3173574.3173677 pp. 1–13.

[576] S. Gregor and I. Benbasat, "Explanations from Intelligent Systems: Theoretical Foundations and Implications for Practice," *MIS Quarterly*, vol. 23, no. 4, pp. 497–530, 1999. DOI: 10.2307/249487

[577] S. Anjomshoae, K. Främling, and A. Najjar, "Explanations of Black-Box Model Predictions by Contextual Importance and Utility," in *Explainable, Transparent Autonomous Agents and Multi-Agent Systems*. Springer, 2019. DOI: 10.1007/978-3-030-30391-4_6 pp. 95–109.

[578] F. Hohman, A. Head, R. Caruana, R. DeLine, and S. M. Drucker, "Gamut: A design probe to understand how data scientists understand machine learning models," in *Proceedings of the 2019 Conference on Human Factors in Computing Systems (CHI)*. New York, NY, USA: Association for Computing Machinery, 2019. DOI: 10.1145/3290605.3300809 pp. 1–13.

[579] L. A. Hendricks, Z. Akata, M. Rohrbach, J. Donahue, B. Schiele, and T. Darrell, "Generating visual explanations," in *European Conference on Computer Vision*. Springer, 2016, pp. 3–19.

[580] R. Sevastjanova, F. Beck, B. Ell, C. Turkay, R. Henkin, M. Butt, D. A. Keim, and M. El-Assady, "Going beyond Visualization: Verbalization as Complementary Medium to Explain Machine Learning Models," in *VIS Workshop on Visualization for AI Explainability (VISxAI)*, 2018, pp. 1–6.

[581] S. Sreedharan, T. Chakraborti, and S. Kambhampati, "Handling Model Uncertainty and Multiplicity in Explanations via Model Reconciliation," *Proceedings of the International Conference on Automated Planning and Scheduling*, vol. 28, no. 1, 2018.

[582] M. M. De Graaf and B. F. Malle, "How people explain action (and autonomous intelligent systems should too)," in *2017 AAAI Fall Symposium Series*, 2017.

[583] M. O. Riedl, "Human-centered artificial intelligence and machine learning," *Human Behavior and Emerging Technologies*, vol. 1, no. 1, pp. 33–36, 2019. DOI: 10.1002/hbe2.117

[584] O. Biran and K. McKeown, "Human-Centric Justification of Machine Learning Predictions," in *Proceedings of the 26th International Joint Conference on Artificial Intelligence*. AAAI Press, 2017, p. 1461–1467.

[585] B. Lamche, U. Adıgüzel, and W. Wörndl, "Interactive explanations in mobile shopping recommender systems," in *Joint Workshop on Interfaces and Human Decision Making in Recommender Systems*, vol. 14, 2014.

[586] J. Chen, F. Lecue, J. Z. Pan, I. Horrocks, and H. Chen, "'Knowledge-Based Transfer Learning Explanation," in *Sixteenth International Conference on Principles of Knowledge Representation and Reasoning*, 2018, pp. 349–358.

[587] B. P. Knijnenburg and A. Kobsa, "Making decisions about privacy: information disclosure in context-aware recommender systems," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 3, no. 3, pp. 1–23, 2013.

[588] J. Zhou, H. Hu, Z. Li, K. Yu, and F. Chen, "Physiological Indicators for User Trust in Machine Learning with Influence Enhanced Fact-Checking," in *International Cross-Domain Conference for Machine Learning and Knowledge Extraction.* Springer, 2019. DOI: 10.1007/978-3-030-29726-8_7 pp. 94–113.

[589] M. K. Lee, A. Jain, H. J. Cha, S. Ojha, and D. Kusbit, "Procedural Justice in Algorithmic Fairness," *Proceedings of the 2019 ACM on Human-Computer Interaction*, vol. 3, pp. 1–26, 2019. DOI: 10.1145/3359284

[590] M. J. Druzdzel, "Qualitative verbal explanations in bayesian belief networks," *AISB QUARTERLY*, pp. 43–54, 1996.

[591] U. Ehsan, B. Harrison, L. Chan, and M. O. Riedl, "Rationalization: A neural machine translation approach to generating natural language explanations," in *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, 2018, pp. 81–87.

[592] M. R. Wick and W. B. Thompson, "Reconstructive expert system explanation," *Artificial Intelligence*, vol. 54, no. 1-2, pp. 33–70, 1992.

[593] J. Van Bouwel and E. Weber, "Remote causes, bad explanations?" *Journal for the Theory of Social Behaviour*, vol. 32, no. 4, pp. 437–449, 2002.

[594] O. Zinovatna and L. M. Cysneiros, "Reusing knowledge on delivering privacy and transparency together," in *2015 IEEE fifth international workshop on requirements patterns (RePa).* IEEE, 2015, pp. 17–24.

[595] J. F. Osborne and A. Patterson, "Scientific argument and explanation: A necessary distinction?" *Science Education*, vol. 95, no. 4, pp. 627–638, 2011.

[596] J. D. Trout, "Scientific explanation and the sense of understanding," *Philosophy of Science*, vol. 69, no. 2, pp. 212–233, 2002.

[597] L. Edwards and M. Veale, "Slave to the algorithm: Why a right to an explanation is probably not the remedy you are looking for," *Duke L. & Tech. Rev.*, vol. 16, p. 18, 2017.

[598] I. Baaj, J.-P. Poli, and W. Ouerdane, "Some Insights Towards a Unified Semantic Representation of Explanation for eXplainable Artificial Intelligence," in *Proceedings of the 2019 Workshop on Interactive Natural Language Technology for Explainable Artificial Intelligence*

(NL4XAI). Association for Computational Linguistics, 2019. DOI: 10.18653/v1/W19-8404 pp. 14–19.

[599] M. Strevens, "The causal and unification approaches to explanation unified—causally," *Noûs*, vol. 38, no. 1, pp. 154–176, 2004.

[600] F. Nothdurft, T. Heinroth, and W. Minker, "The Impact of Explanation Dialogues on Human-Computer Trust," in *International Conference on Human-Computer Interaction (HCI)*. Springer, 2013. DOI: 10.1007/978-3-642-39265-8_7 pp. 59–67.

[601] L. R. Ye and P. E. Johnson, "The impact of explanation facilities on user acceptance of expert systems advice," *Mis Quarterly*, pp. 157–172, 1995.

[602] A. Vellido, "The importance of interpretability and visualization in machine learning for applications in medicine and health care," *Neural computing and applications*, pp. 1–15, 2019.

[603] T. Lombrozo, "The Instrumental Value of Explanations," *Philosophy Compass*, vol. 6, no. 8, pp. 539–551, 2011. DOI: 10.1111/j.1747-9991.2011.00413.x

[604] J. Trout, "The psychology of scientific explanation," *Philosophy Compass*, vol. 2, no. 3, pp. 564–591, 2007.

[605] K. McCarthy, J. Reilly, L. McGinty, and B. Smyth, "Thinking positively-explanatory feedback for conversational recommender systems," in *Proceedings of the European Conference on Case-Based Reasoning (ECCBR-04) Explanation Workshop*, 2004, pp. 115–124.

[606] T.-W. Chen and S. S. Sundar, "This App Would Like to Use Your Current Location to Better Serve You: Importance of User Assent and System Transparency in Personalized Mobile Services," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI '18. New York, NY, USA: Association for Computing Machinery, 2018. DOI: 10.1145/3173574.3174111 p. 1–13.

[607] M. Sridharan and B. Meadows, "Towards a Theory of Explanations for Human – Robot Collaboration," *KI-Künstliche Intelligenz*, vol. 33, no. 4, pp. 331–342, 2019. DOI: 10.1007/s13218-019-00616-y

[608] V. Dominguez, P. Messina, C. Trattner, and D. Parra, "Towards Explanations for Visual Recommender Systems of Artistic Images," in *Joint Workshop on Interfaces and Human Decision Making for Recommender Systems* , 2018.

[609] M. Atzmueller, "Towards Socio-Technical Design of Explicative Systems: Transparent, Interpretable and Explainable Analytics and Its Perspectives in Social Interaction Contexts information," in *Proceedings of the 2019 Workshop on Affective Computing and Context Awareness in Ambient Intelligence (AfCAI)*, 2019, pp. 1–8.

[610] R. Iyer, Y. Li, H. Li, M. Lewis, R. Sundar, and K. Sycara, "Transparency and Explanation in Deep Reinforcement Learning Neural Networks," in *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, ser. AIES '18. New York, NY, USA: Association for Computing Machinery, 2018. DOI: 10.1145/3278721.3278776 p. 144–150.

[611] K. Balog, F. Radlinski, and S. Arakelyan, "Transparent, Scrutable and Explainable User Models for Personalized Recommendation," in *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval.* New York, NY, USA: Association for Computing Machinery, 2019. DOI: 10.1145/3331184.3331211 pp. 265–274.

[612] A. Abdul, J. Vermeulen, D. Wang, B. Y. Lim, and M. Kankanhalli, "Trends and Trajectories for Explainable, Accountable and Intelligible Systems: An HCI Research Agenda," in *Proceedings of the 2018 Conference on Human Factors in Computing Systems (CHI).* New York, NY, USA: Association for Computing Machinery, 2018. DOI: 10.1145/3173574.3174156 pp. 1–18.

[613] M. A. Neerincx, J. van der Waa, F. Kaptein, and J. van Diggelen, "Using perceptual and cognitive explanations for enhanced human-agent team performance," in *International Conference on Engineering Psychology and Cognitive Ergonomics.* Springer, 2018, pp. 204–214.

[614] J. Bidot, S. Biundo-Stephan, T. Heinroth, W. Minker, F. Nothdurft, and B. Schattenberg, "Verbal Plan Explanations for Hybrid Planning," in *MKWI*, 2010.

[615] S. Rosenthal, S. P. Selvaraj, and M. Veloso, "Verbalization: Narration of Autonomous Robot Experience," in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, ser. IJCAI'16. AAAI Press, 2016, p. 862–868.

[616] D. Doran, S. Schulz, and T. R. Besold, "What Does Explainable AI Really Mean? A New Conceptualization of Perspectives," in *Proceedings of the First International Workshop on Comprehensibility and Explanation in AI and ML (CEX 2017)*, 2017.

[617] R. Sheh, ""Why did you do that?Ëxplainable intelligent robots," in *AAAI Workshop-Technical Report*, 2017, pp. 628–634.

[618] T. Kulesza, S. Stumpf, W.-K. Wong, M. M. Burnett, S. Perona, A. Ko, and I. Oberst, "Why-oriented end-user debugging of naive Bayes text classification," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 1, no. 1, pp. 1–31, 2011. DOI: 10.1145/2030365.2030367

[619] L. Chazette, W. Brunotte, and T. Speith, "Supplementary Material for Research Paper „Exploring Explainability: A Definition, a Model, and a Knowledge Catalogue"," 2021. DOI: 10.5281/zenodo.5114922

[620] M. Hagen, "User-Centered Privacy Communication Design," in *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*. Denver, CO: USENIX Association, 2016.

[621] S. Korff and R. Böhme, "Too Much Choice: End-User Privacy Decisions in the Context of Choice Proliferation," in *10th Symposium On Usable Privacy and Security (SOUPS 2014)*. Menlo Park, CA: USENIX Association, 2014, pp. 69–87.

[622] B. Abdollahi and O. Nasraoui, "Transparency in fair machine learning: the case of explainable recommender systems," in *Human and machine learning*. Springer, 2018, pp. 21–35.

[623] M. Langer, K. Baum, K. Hartmann, S. Hessel, T. Speith, and J. Wahl, "Explainability Auditing for Intelligent Systems: A Rationale for Multi-Disciplinary Perspectives," in *2021 IEEE 29th International Requirements Engineering Conference Workshops (REW)*. Ieee, 2021, pp. 164–168.

[624] J. Schneider and J. Handali, "Personalized explanation in machine learning: A conceptualization," *arXiv preprint arXiv:1901.00770*, 2019. DOI: 10.48550/arXiv.1901.00770

[625] S. Thiebes, S. Lins, and A. Sunyaev, "Trustworthy artificial intelligence," *Electronic Markets*, vol. 31, no. 2, pp. 447–464, 2021.

[626] D. Ameller, C. Ayala, J. Cabot, and X. Franch, "How do software architects consider non-functional requirements: An exploratory study," in *2012 20th IEEE international requirements engineering conference (RE)*. Ieee, 2012, pp. 41–50.

[627] B. Boehm and H. In, "Identifying quality-requirement conflicts," *IEEE software*, vol. 13, no. 2, pp. 25–35, 1996.

[628] T. Breaux, "Towards Regulatory Compliance: Extracting Rights and Obligations to Align Requirements with Regulations," in *Proc. of 14th IEEE International Requirements Engineering Conference, 2006*, 2006.

[629] Centre for International Governance Innovation (CIGI), "CIGI-Ipsos Global Survey on Internet Security and Trust 2019," https://www.cigionline.org/cigi-ipsos-global-survey-internet-security-and-trust/, 2019, last visited on 2022-04-14.

[630] L. Chung and J. C. S. d. Prado Leite, "On non-functional requirements in software engineering," in *Conceptual modeling: Foundations and applications*. Springer, 2009, pp. 363–379.

[631] L. Coles-Kemp, R. B. Jensen, and C. P. Heath, "Too much information: questioning security in a post-digital society," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–14.

[632] C. L. Corritore, B. Kracher, and S. Wiedenbeck, "On-line trust: concepts, evolving themes, a model," *International journal of human-computer studies*, vol. 58, no. 6, pp. 737–758, 2003.

[633] L. F. Cranor and S. Garfinkel, "Guest Editors' Introduction: Secure or Usable?" *IEEE security & privacy*, vol. 2, no. 5, pp. 16–18, 2004.

[634] J. L. De La Vara, K. Wnuk, R. Berntsson-Svensson, J. Sánchez, and B. Regnell, "An Empirical Study on the Importance of Quality Requirements in Industry." in *Seke*, 2011, pp. 438–443.

[635] V. Distler, M.-L. Zollinger, C. Lallemand, P. Roenne, P. Ryan, and V. Koenig, "Security–visible, yet unseen? how displaying security mechanisms impacts user experience and perceived security," in *Proceedings of ACM CHI Conference on Human Factors in Computing Systems (CHI2019)*, 2019.

[636] C. Flavián, M. Guinalíu, and R. Gurrea, "The role played by perceived usability, satisfaction and consumer trust on website loyalty," *Information & management*, vol. 43, no. 1, pp. 1–14, 2006.

[637] E. C. Groen, S. Kopczyńska, M. P. Hauer, T. D. Krafft, and J. Doerr, "Users – the hidden software product quality experts?: A study on how app users report quality aspects in online reviews," in *2017 IEEE 25th international requirements engineering conference (RE)*. Ieee, 2017, pp. 80–89.

[638] P. Gutmann and I. Grigg, "Security usability," *IEEE security & privacy*, vol. 3, no. 4, pp. 56–58, 2005.

[639] R. Hardin, *Trust and Trustworthiness*. New York, USA: Russell Sage Foundation, 2002.

[640] K. Hawley, *How to be Trustworthy*. Oxford University Press, USA, 2019.

[641] A. Jacovi, A. Marasović, T. Miller, and Y. Goldberg, "Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in AI," in *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, 2021, pp. 624–635.

[642] D. E. Kieras and S. Bovair, "The role of a mental model in learning to operate a device," *Cognitive science*, vol. 8, no. 3, pp. 255–273, 1984.

[643] J. C. S. d. P. Leite and C. Cappelli, "Software transparency," *Business & Information Systems Engineering*, vol. 2, no. 3, pp. 127–139, 2010.

[644] A. F. Markus, J. A. Kors, and P. R. Rijnbeek, "The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies," *Journal of Biomedical Informatics*, vol. 113, p. 103655, 2021.

[645] J. Riegelsberger, M. A. Sasse, and J. D. McCarthy, "The mechanics of trust: A framework for research and design," *International Journal of Human-Computer Studies*, vol. 62, no. 3, pp. 381–422, 2005.

[646] K. Vaccaro, D. Huang, M. Eslami, C. Sandvig, K. Hamilton, and K. Karahalios, "The illusion of control: Placebo effects of control settings," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–13.

[647] K.-P. Yee, "Aligning security and usability," *IEEE Security & Privacy*, vol. 2, no. 5, pp. 48–55, 2004.

[648] E. N. Zalta, U. Nodelman, and C. Allen, *Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Center for the Study of Language and Information, 1995.

[649] B. M. Napoleão, K. R. Felizardo, E. F. d. Souza, F. Petrillo, S. Hallé, N. L. Vijaykumar, and E. Y. Nakagawa, "Establishing a Search String to Detect Secondary Studies in Software Engineering," in *2021 47th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, 2021. DOI: 10.1109/SEAA53835.2021.00010 pp. 9–16.

[650] I. Shklovski, S. D. Mainwaring, H. H. Skúladóttir, and H. Borgthorsson, "Leakiness and Creepiness in App Space: Perceptions of Privacy and Mobile App Use," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '14. New York, NY, USA: Association for Computing Machinery, 2014. DOI: 10.1145/2556288.2557421 p. 2347–2356.

[651] H. Almuhimedi, F. Schaub, N. Sadeh, I. Adjerid, A. Acquisti, J. Gluck, L. F. Cranor, and Y. Agarwal, "Your location has been shared 5,398 times! A field study on mobile app privacy nudging," in *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, 2015, pp. 787–796.

[652] G. Bal, "Designing privacy indicators for smartphone app markets: A new perspective on the nature of privacy risks of apps," in *Proceedings of the 20th Americas Conference on Information Systems*, 2014.

[653] R. Balebako, A. Marsh, J. Lin, J. I. Hong, and L. F. Cranor, "The privacy and security behaviors of smartphone app developers," *Usec '14*, 2014.

[654] Z. Belkhamza, M. Niasin, and A. Faris, "The Effect of Privacy Concerns on Smartphone App Purchase in Malaysia: Extending the Theory of Planned Behavior." *International Journal of Interactive Mobile Technologies*, vol. 11, no. 5, 2017.

[655] K. Bongard-Blanchy, A. Rossi, S. Rivas, S. Doublet, V. Koenig, and G. Lenzini, ""I am Definitely Manipulated, Even When I am Aware of it. It's Ridiculous!"-Dark Patterns from the End-User Perspective," in *Designing Interactive Systems Conference 2021*, 2021, pp. 763–776.

[656] C. Buck, C. Horbel, and T. Eymann, "Dealing with Privacy and Security Risks: App Consumers in Mobile Ecosystems," *Tagungsband Multikonferenz Wirtschaftsinformatik 2014 (MKWI 2014)*, pp. 64–74, 2014.

[657] R. Calo, "Against notice skepticism in privacy (and elsewhere)," *Notre Dame L. Rev.*, vol. 87, p. 1027, 2011.

[658] C. Castelluccia, S. Guerses, M. Hansen, J. Hoepman, J. van Hoboken, B. Vieira, and European Union Agency for Cybersecurity, *Privacy and data protection in mobile applications: A study on the app development ecosystem and the technical implementation of GDPR*. ENISA, the European Union Agency for Network and Information Security, 2018.

[659] I. Chong, H. Ge, N. Li, and R. W. Proctor, "Influence of privacy priming and security framing on mobile app selection," *Computers & Security*, vol. 78, pp. 143–154, 2018.

[660] F. Ebrahimi, M. Tushev, and A. Mahmoud, "Mobile app privacy in software engineering research: A systematic mapping study," *Information and Software Technology*, vol. 133, p. 106466, 2021.

[661] N. Good, R. Dhamija, J. Grossklags, D. Thaw, S. Aronowitz, D. Mulligan, and J. Konstan, "Stopping spyware at the gate: a user study of privacy, notice and spyware," in *Proceedings of the 2005 symposium on Usable privacy and security*, 2005, pp. 43–52.

[662] C. M. Gray, J. Chen, S. S. Chivukula, and L. Qu, "End User Accounts of Dark Patterns as Felt Manipulation," *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. Cscw2, pp. 1–25, 2021.

[663] M. Hatamian, "Engineering privacy in smartphone apps: A technical guideline catalog for app developers," *IEEE Access*, vol. 8, pp. 35 429–35 445, 2020.

[664] M. Hintze, "In defense of the long privacy statement," *Md. L. Rev.*, vol. 76, p. 1044, 2016.

[665] M. J. Jafar, A. Abdullat *et al.*, "Exploratory analysis of the readability of information privacy statement of the primary social networks," *Journal of Business & Economics Research (JBER)*, vol. 7, no. 12, 2009.

[666] S. Joeckel and L. Dogruel, "Default effects in app selection: German adolescents' tendency to adhere to privacy or social relatedness features in smartphone apps," *Mobile Media & Communication*, vol. 8, no. 1, pp. 22–41, 2020.

[667] R. Kesler, M. Kummer, and P. Schulte, "Competition and privacy in online markets: Evidence from the mobile app industry," in *Academy of Management Proceedings*. Academy of Management Briarcliff Manor, NY 10510, 2020, p. 20978.

[668] O. Kulyk, P. Gerber, M. El Hanafi, B. Reinheimer, K. Renaud, and M. Volkamer, "Encouraging Privacy-Aware Smartphone App Installation: What Would the Technically-Adept Do," in *USEC'16-Usable Security Workshop, 21 February 2016, San Diego*. Internet Society, 2016.

[669] B. Liu, J. Lin, and N. Sadeh, "Reconciling mobile app privacy and usability on smartphones: Could user privacy profiles help?" in *Proceedings of the 23rd international conference on World wide web*, 2014, pp. 201–212.

[670] A. M. McDonald and T. Lowenthal, "Nano-notice: Privacy disclosure at a mobile scale," *Journal of Information Policy*, vol. 3, no. 1, pp. 331–354, 2013.

[671] A. N. Mehdy and H. Mehrpouyan, "Modeling of personalized privacy disclosure behavior: A formal method approach," in *The 16th International Conference on Availability, Reliability and Security*, 2021, pp. 1–13.

[672] G. R. Milne, M. J. Culnan, and H. Greene, "A longitudinal assessment of online privacy notice readability," *Journal of Public Policy & Marketing*, vol. 25, no. 2, pp. 238–249, 2006.

[673] F. Mosca, "Value-Aligned and Explainable Agents for Collective Decision Making: Privacy Application." in *Aamas*, 2020, pp. 2199–2200.

[674] J. R. Reidenberg, N. C. Russell, A. J. Callen, S. Qasir, and T. B. Norton, "Privacy harms and the effectiveness of the notice and choice framework," *Isjlp*, vol. 11, p. 485, 2015.

[675] H. Zhu, H. Xiong, Y. Ge, and E. Chen, "Mobile app recommendations with security and privacy awareness," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2014, pp. 951–960.

Liste der wissenschaftlichen Publikationen

[1] W. Brunotte, A. Specht, L. Chazette, and K. Schneider, "Privacy Explanations – A Means to End-User Trust," *Journal of Systems and Software*, vol. 195, p. 111545, 2023. DOI: 10.1016/j.jss.2022.111545

[2] W. Brunotte, J. Droste, and K. Schneider, "Context, Content, Consent – How to Design User-Centered Privacy Explanations," in *The 35th International Conference on Software Engineering & Knowledge Engineering*, San Francisco, USA, 2023. DOI: 10.18293/SEKE2023-032. ISSN 2325-9000 pp. 86–89.

[3] W. Brunotte and J. Droste, "Supplementary Material for Research Paper „Context, Content, Consent - How to Design User- Centered Privacy Explanations"," Mai 2023. DOI: 10.5281/zenodo.7911904

[4] W. Brunotte, L. Chazette, L. Köhler, J. Klunder, and K. Schneider, "What About My Privacy? Helping Users Understand Online Privacy Policies," in *Proceedings of the International Conference on Software and System Processes and International Conference on Global Software Engineering*, ser. Icssp'22. New York, NY, USA: Association for Computing Machinery, 2022. DOI: 10.1145/3529320.3529327 p. 56–65.

[5] W. Brunotte, L. Chazette, V. Klös, and T. Speith, "Quo Vadis, Explainability? – A Research Roadmap for Explainability Engineering," in *Requirements Engineering: Foundation for Software Quality*, V. Gervasi and A. Vogelsang, Eds. Cham: Springer International Publishing, 2022, pp. 26–32.

[6] L. Chazette, W. Brunotte, and T. Speith, "Explainable Software Systems: From Requirements Analysis to System Evaluation," *Requirements Engineering*, vol. 27, no. 4, pp. 457–487, Dec 2022. DOI: 10.1007/s00766-022-00393-5

[7] W. Brunotte, L. Nagel, K. Schneider, and J. Klünder, "How to Identify Changing Contexts of Use with Creativity Workshops – An Experience Report," in *Sense, Feel, Design*,

C. Ardito, R. Lanzilotti, A. Malizia, M. Larusdottir, L. D. Spano, J. Campos, M. Hertzum, T. Mentler, J. Abdelnour Nocera, L. Piccolo, S. Sauer, and G. van der Veer, Eds. Cham: Springer International Publishing, 2022. DOI: 10.1007/978-3-030-98388-8_9 pp. 88–97.

[8] W. Brunotte, L. Chazette, V. Klös, and T. Speith, "Supplementary Material for Vision Paper „Quo Vadis, Explainability? – A Research Roadmap for Explainability Engineering"," Januar 2022. DOI: 10.5281/zenodo.5902181

[9] W. Brunotte, "Data for Research Article „Privacy Explanations – A Means to End-User Trust"," October 2022. DOI: 10.5281/zenodo.7215560

[10] W. Brunotte, L. Chazette, and K. Korte, "Can Explanations Support Privacy Awareness? A Research Roadmap," in *2021 IEEE 29th International Requirements Engineering Conference Workshops (REW)*, 2021. DOI: 10.1109/rew53955.2021.00032 pp. 176–180.

[11] 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)*, 2021. DOI: 10.1109/RE51729.2021.00025 pp. 197–208.

[12] W. Brunotte, L. Chazette, V. Klos, E. Knauss, T. Speith, and A. Vogelsang, "Welcome to the First International Workshop on Requirements Engineering for Explainable Systems (RE4ES)," in *2021 IEEE 29th International Requirements Engineering Conference Workshops (REW)*. Los Alamitos, CA, USA: IEEE Computer Society, sep 2021. DOI: 10.1109/REW53955.2021.00028 pp. 157–158.

[13] F. P. Viertel, W. Brunotte, Y. Evers, and K. Schneider, "Community Knowledge About Security:," in *Risks and Security of Internet and Systems*, J. Garcia-Alfaro, J. Leneutre, N. Cuppens, and R. Yaich, Eds. Cham: Springer International Publishing, 2021, pp. 181–197.

[14] F. Kortum, J. Klünder, O. Karras, W. Brunotte, and K. Schneider, "Which Information Help agile Teams the Most? An Experience Report on the Problems and Needs," in *2020 46th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, 2020. DOI: 10.1109/SEAA51224.2020.00058 pp. 306–313.

[15] J. Klünder, W. Brunotte, and K. Schneider, "When you don't know with whom to collaborate: Towards an interactive system connecting contributors in a research project," in *HCI Engineering 2019: HCI Engineering 2019 - methods and tools for advanced interactive systems and integration of multiple stakeholder viewpoints : joint proceedings HCI Engineering 2019 - methods and tools for advanced interactive systems and integration of*

multiple stakeholder viewpoints, co-located with 11th ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS 2019). Aachen, Germany : RWTH Aachen, 2019 (CEUR workshop proceedings, B. Weyers and J. Bowen, Eds. Valencia, Spain: CEUR workshop, 2019. DOI: 10.15488/9403 pp. 1–8.

[16] F. Kortum, J. Klünder, W. Brunotte, and K. Schneider, "Sprint Performance Forecasts in Agile Software Development: The Effect of Futurespectives on Team-Driven Dynamics," in *The 31th International Conference on Software Engineering & Knowledge Engineering*, Lissabon, Portugal, 2019. DOI: 10.18293/SEKE2019-224 pp. 94–101.

[17] P. Viertel, Fabien, W. Brunotte, D. Strüber, and K. Schneider, "Detecting Security Vulnerabilities using Clone Detection and Community Knowledge," in *The 31th International Conference on Software Engineering & Knowledge Engineering*, Lissabon, Portugal, 2019. DOI: 10.18293/SEKE2019-183 pp. 94–101.

[18] W. Brunotte, "Security Code Clone Detection entwickelt als Eclipse Plugin," Master's thesis, Leibniz Universität Hannover, Fachgebiet Software Engineering, Hannover, Germany, June 2018.

[19] W. Brunotte, "On-the-fly Synthese für szenariobasierte Spezifikationen von Produktlinien," Bachelor, Leibniz Universität Hannover, Fachgebiet Software Engineering, März 2015.