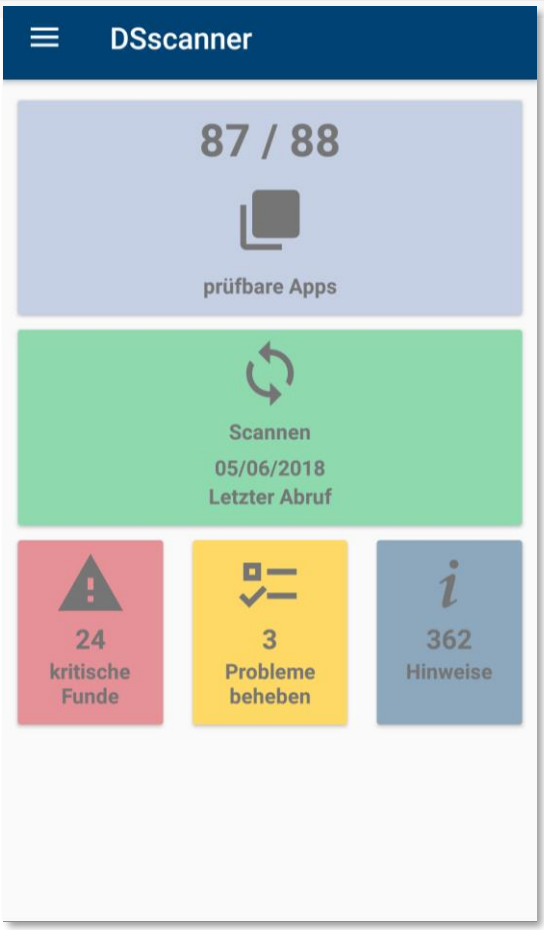
Gefördert vom

Bundesministerium
für Bildung
und Forschung

PGuard

Gemeinsamer Abschlussbericht

Dr. Sara Elisa Kettner, Prof. Dr. Christian Thorun
Sascha Ludwig M.Sc., Prof. Dr. Gerhard Heyer
Dipl.-Des. Wulf Bolte, Sebastian Wolters B.A.
Frank Ingenrieth LL.M., Carolin Rost, Jörn Wittmann



selbstregulierung
informationswirtschaft e.V.

mediaTest
digital

InfAI
Institut für Angewandte Informatik

Quadriga
Hochschule

Dies ist der gemeinsame Abschlussbericht des Forschungskonsortiums „PGuard“

- Institut für Angewandte Informatik e. V., Goerdelerring 9, 04109 Leipzig
- mediaTest digital GmbH, Goseriede 4, 30159 Hannover
- Quadriga Hochschule Berlin GmbH, Werderscher Markt 13, 10117 Berlin
- Selbstregulierung Informationswirtschaft e.V., Albrechtsstraße 10B, 10117 Berlin

https://datenschutz-scanner.de

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung gefördert (koordinierendes Förderkennzeichen: 16KIS0388K).

Die Verantwortung für den Inhalt dieses Berichts liegt bei den Autoren.

Zitiervorschlag:

Abschlussbericht PGuard, Kettner/Bolte/Heyer/Ingenrieth/Ludwig/Thorun u.a., 2019.

Inhaltsverzeichnis

1 Förderpolitische Ziele des Forschungsprojektes

Das Forschungsprojekt wurde im Rahmen der Bekanntmachung „Datenschutz: selbstbestimmt in der digitalen Welt“[1] durchgeführt. Die unter dieser Bekanntmachung durchgeführten Forschungsprojekte sollten ein durch den Zuwendungsgeber ermitteltes Defizit an „wirksame[n] und alltagstaugliche[n] Ansätze[n] für den Selbstdatenschutz in der digitalen Welt“[2] ausräumen. Dabei hat sich dieses Forschungsvorhaben den Bedarfen „hinsichtlich innovativer und alltagstauglicher Lösungsansätze, Technologien und komplementärer Maßnahmen, welche Laien auch im privaten Kontext erst befähigen, die Datenschutzrisiken besser einzuschätzen“ gewidmet. Hierbei wurde sich darauf konzentriert, dass die „Weitergabe und Nutzung von Daten […] verständlich und nachvollziehbar dargestellt werden“[3]. Der Fokus des Forschungsvorhabens auf mobile Smartphone-Apps adressiert hierbei den im Rahmen der Bekanntmachung als wichtig hervorgehobenen mobilen Anwendungszusammenhang, da „insbesondere die mobile Nutzung des Internets […] heute noch mit einem erhöhten Nutzungsrisiko belastet“[4] sei.

Die Ergebnisse[5] dieses Forschungsprojektes tragen erheblich zu den förderpolitischen Zielen bei. Zunächst wurde, durch die (Weiter-)Entwicklung eines strukturierten und standardisierten Layered-Approachs, in Form der Informationstexte[6], zur Aufbereitung der Informationen für Verbraucherinnen und Verbraucher, der Grundstein für die Nachvollziehbarkeit der Datenverarbeitungen durch Verständnis gelegt. Dieser Ansatz wurde zudem mit einer für Verbraucherinnen und Verbraucher verständlichen Sprache umgesetzt. Dank des Layered-Approachs können– ohne die Verständlichkeit zu verlieren – auch Informationen transportiert und vermittelt werden, die eine Risikoeinschätzung der jeweiligen Datenverarbeitungen ermöglicht. Eine solche Risikobewertung wird durch den im Forschungsprojekt entwickelten Ansatz in doppelter Hinsicht adressiert: einerseits wird für jede Form der Datenverarbeitung transparent über den Umstand sowie damit einhergehende Vor- und Nachteile informiert. Aufgrund dieser Informationen sind Verbraucherinnen und Verbraucher in der Lage, selbstbestimmt eine Entscheidung bezüglich der konkreten Datenverarbeitung zu treffen. Andererseits wurden Datenverarbeitungsumstände definiert, die zu einer gesonderten Warnung führen. Diese sogenannten „roten Linien“[7] bilden zwar nur einen kleinen Anteil der aufbereiteten Aspekte ab, stellen aber dafür kritische Umstände bzw. Gesetzesverstöße dar. Die Selbstbestimmung der Verbraucherinnen und Verbraucher soll in diesem Kontext nicht so stark reduziert werden, zumal bestimmte – formelle Gesetzverstöße – nicht unbedingt mit einem tatsächlich schlechteren Schutzniveau einhergehen müssen. Insofern können auch in diesen Fällen Verbraucherinnen und Verbraucher entsprechende Hinweise als „für sich selbst unproblematisch“ einstufen.

[1] https://www.bmbf.de/foerderungen/bekanntmachung.php?B=971.
[2] https://www.bmbf.de/foerderungen/bekanntmachung.php?B=971.
[3] https://www.bmbf.de/foerderungen/bekanntmachung.php?B=971.
[4] https://www.bmbf.de/foerderungen/bekanntmachung.php?B=971.
[5] Zusammenfassung der wesentlichen Ergebnisse in Kapitel 2.
[6] Siehe 4.8.2.
[7] Siehe 4.8.2.2.

Diese erhöhte Nachvollziehbarkeit basiert auf zwei Informationsquellen, nämlich der technischen und der semantischen Analyse. Hierdurch entsteht ein sehr umfangreiches Abbild der möglichen (und tatsächlichen) Datenverarbeitungen. Dieses Abbild ist in der Lage ein erhöhtes Vertrauen in die Datenverarbeitungen zu fördern, da Datenverarbeitungen, die Verbraucherinnen und Verbraucher ansonsten nicht zur Kenntnis genommen hätten, durch die im Forschungsprojekt entwickelten Demonstratoren leicht zugänglich und transparent zur Kenntnis genommen werden können.

Die technische Analyse trägt zudem auch im Besonderen zu einer erhöhten Vertraulichkeit bei. Eine fehlende oder unzureichende Verschlüsselung bei der Datenübermittlung kann festgestellt und den Verbraucherinnen und Verbrauchern kommuniziert werden. Dies ist im App-Kontext wichtig, da dort Verbraucherinnen und Verbraucher in der Regel nicht in der Lage sind, entsprechende Prüfungen selbst vorzunehmen; eine standardisierte „Kommunikationsform“, wie diese für Webseiten in Form von HTTPS und „grünen“ bzw. „roten Schlössern“ in der Adresszeile besteht, fehlt für Datenströme durch mobile Applikationen.

Die gewählten Umsetzungsformen im Rahmen des Forschungsprojektes haben auch die ethischen, rechtlichen, sozialwissenschaftlichen und wirtschaftswissenschaftlichen Implikationen berücksichtigt, ohne dabei die Selbstbestimmung der Verbraucherinnen und Verbraucher aus den Augen zu verlieren. So stellt die im Forschungsprojekt entwickelte Lösung keinen „Pranger“ im klassischen Sinne dar, in der ausschließlich negative Aspekte hervorgehoben werden. Vielmehr werden – auch in Rückkopplung mit Verbraucherinnen und Verbraucher durch Umfragen - Aspekte kommuniziert, die für Verbraucherinnen und Verbraucher von Interesse sind und dies kontextbezogen, verständlich und informativ. Da Verbraucherinnen und Verbraucher zudem konkrete Handlungsempfehlungen erhalten, die soweit möglich eine Nutzung betroffener Apps weiterhin ermöglicht, ist dennoch von einem möglichen, dauerhaften Marktdruck auf App-Anbieter auszugehen, welcher im Ergebnis zu transparenteren Datenverarbeitungen führen könnte und hierbei insbesondere „schwarze Schafe“ verdrängt.

2 Zusammenfassung und wesentliche Ergebnisse

Ziel des Forschungsprojektes war es den Selbstdatenschutz der Bürgerinnen und Bürger zu stärken und Lösungen zur Ausübung der informationellen Selbstbestimmung zu entwickeln. Es wurden diesbezüglich mehrere konzeptionelle und empirische Teilergebnisse erzielt, die zur Entwicklung des Endresultats in Form dreier Labormuster notwendig waren:

- Untersuchung der Verbrauchermotivation und -kenntnisse
- Entwicklung verständlicher Datenschutzinformationen, die Verbraucherinnen und Verbrauchern den Zugang zu Datenschutztexten erleichtern
- Erarbeitung eines Nutzerführungskonzepts das die informationelle Selbstbestimmung von Verbraucherinnen und Verbrauchern fördert
- Automatisierte Suche sowie Übertragung von Datenschutzerklärungen in Appstores und innerhalb von Apps
- Automatisierter Abgleich unterschiedlicher Versionen von Datenschutzerklärungen
- Automatisierte Analyse von Datenschutzerklärungstexten hinsichtlich spezifischer Inhalte
- Automatisierter Abruf von Installationsdateien in der jeweils notwendigen Version
- Automatisierte Aufbereitung und Analyse von Installationsdateien hinsichtlich spezifischer Inhalte
- Kommunikations- und Programmschnittstellen zum Austausch bzw. Abruf von Prüfergebnissen sowie Anstoßen weiterer Prüfungen
- Analyse und datenschutzfreundliche Umsetzung von betriebssystemspezifischen Funktionen – Auslesen installierter Apps, Verlinkung in Einstellungen und Funktionen des mobilen Endgerätes etc.

Aus den Teilergebnissen leiteten sich drei Labormuster ab, namentlich der DATENSCHUTZscanner App-Client, der „Check Your-APPS[8]“ Datenschutzerklärungs-Analyzer und das PGuard Browser-Plugin. Diese ermöglichen Verbraucherinnen und Verbrauchern in ihrer Anwendung den selbstbestimmten Umgang mit ihren personenbezogenen Daten.

Darüber hinaus wurden im Rahmen des Projekts verschiedene Optionen untersucht, die die Verwertung der Labormuster ermöglichen könnten.

2.1 Methodik

Zu den Methoden, die zur Erreichung der Teilergebnisse und zur Entwicklung der Labormuster notwendig waren, zählen zwei repräsentative Online Befragungen, drei Fokusgruppengespräche, UX-Tests mit potentiellen Nutzerinnen und Nutzern, sechs Experteninterviews im Rahmen einer Experten Beta-Phase, die Recherche und Umsetzung etablierter technischer und semantischer Analysemethoden, Verifikation und Benchmarking der Testergebnisse je Methode und entsprechende Weiterentwicklung, Recherche

[8] „App Privacy Policy Scanner“

und Analyse der Implikationen durch geänderte rechtliche Rahmenbedingungen, sowie eine systematische Analyse der Wettbewerbsumfelds und sonstiger Marktinteressen. Zusätzlich wurden interaktive Workshop-Formate umgesetzt, die die Leitmotive der Labormuster fortlaufend weiterentwickeln konnten.

2.2 Ergebnisse

Die zentralen Ergebnisse lassen sich wie folgt zusammenfassen:

2.2.1 Ausgangspunkt Verbraucherrealität

Aus den empirischen Erhebungen im Rahmen der Befragungen ergab sich, dass Verbraucherinnen und Verbraucher aktuell einen begrenzten Wissensstand beim Thema Datenschutz und -verarbeitungen haben. Jedoch sind sie grundsätzlich motiviert, Maßnahmen zu ergreifen, die ihren Selbstdatenschutz fördern. So haben Verbraucherinnen und Verbraucher auch ein Interesse an Anwendungen, die sie zu einem selbstbestimmten Handeln befähigen. Funktionen, die im entwickelten Labormuster integriert wurden, werden seitens der Verbraucherinnen und Verbraucher entsprechend nachgefragt.

2.2.2 Bereitstellung von verständlichen Informationen

Um es Verbraucherinnen und Verbrauchern zu ermöglichen, Datenverarbeitungen nachzuvollziehen und gemäß ihren persönlichen Präferenzen handeln zu können, bedarf es verständlicher Informationen zur Datennutzung. Deshalb wurden auf Basis mehrerer Workshops, Fokusgruppengesprächen und Befragungen die sogenannten Informationstexte entwickelt, die auf verständliche und prägnante Weise darstellen, wie personenbezogene Daten verwendet werden. Diese bereiten neben den übergeordneten Beschreibungen der Datenverarbeitungen auch die Vor- und Nachteile auf. Darüber hinaus stellen sie Informationen zu möglichen Handlungsoptionen bereit, die Verbraucherinnen und Verbraucher nutzen können, um ungewollte Datenverarbeitungen zu vermeiden.

2.2.3 Bedienbarkeit der Labormuster

Damit Verbraucherinnen und Verbraucher die zugrundeliegende Technologie anwenden können, bedarf es neben verständlichen Informationen auch eines intuitiven Nutzungskonzepts, das eine einfache Bedienbarkeit und Zugänglichkeit zu den Informationen bereitstellt. Hierzu wurden aus mehreren Nutzer- und Expertengesprächen Bedienkonzepte abgeleitet, die die Bewertung von Datenverarbeitungen gemäß persönlichen Präferenzen ermöglichen. Darüber hinaus beinhalten die Labormuster Vorbewertungen von eindeutigen Gesetzesverstößen und befähigen Verbraucherinnen und Verbraucher die betroffenen Anwendungen zu vermeiden.

2.2.4 Analysetechnologien

Um Nutzerinnen und Nutzern überhaupt verständliche Informationen bereitstellen zu können, mussten im Rahmen des Forschungsprojektes Methoden entwickelt werden, mit deren Hilfe die Informationen zuverlässig ermittelt werden können. Die Informationen werden aus zwei Quellen generiert, der semantischen und der technischen Analyse.

Im Rahmen der **semantischen Analyse** werden die Datenschutzerklärungen der Anbieter analysiert. Hierzu galt es die Datenschutzerklärungen automatisiert zu ermitteln und in ein für die weitere Verarbeitung standardisiertes Format zu übertragen. Hierzu wurde sowohl ein entsprechender Crawler für Appstores entwickelt als auch eine Methode, die es ermöglicht Datenschutzerklärungen innerhalb von Apps zu finden. Im Weiteren wurde eine Annotationslogik entwickelt, die es ermöglicht auf Basis von ca. 300 Einzelangaben vielfältige Erkenntnisse aus einer Datenschutzerklärung abzuleiten. Im Projekt wurde sich auf ca. 30 für Nutzerinnen und Nutzer interessante Erkenntnisse fokussiert, die in Form von Informationstexten[9] aufbereitet wurden. Die zum Einsatz kommenden Logiken sind dynamisch ausgestaltet, sodass einerseits weitere Erkenntnisse mit geringem Aufwand erschlossen werden können und andererseits die Logiken zur Ermittlung der Einzelangaben regelmäßig automatisiert gegen alternative Logiken getestet werden. Letzteres ermöglicht es, dass die Algorithmen zur Analyse der Einzelangaben bei veränderter, beziehungsweise erweiterter, Datengrundlage stets die für die jeweilige Einzelangabe optimale Logik verwenden.

Die **technische Analyse** dient zur Ermittlung der tatsächlich möglichen Datenverarbeitungen einer App. Im Rahmen des Forschungsprojektes wurden diesbezüglich Methoden entwickelt und verfeinert, die entsprechende Aussagen ermöglichen. Ein Großteil der relevanten Informationen können durch eine automatisierte Analyse der Installationsdateien ermittelt werden; ein Rückgriff auf eine manuelle (Datenfluss-) Analyse ist nur im Ausnahmefall erforderlich.

2.2.5 Labormuster insbesondere App-Client

Einerseits galt es die entwickelten Bedienkonzepte umzusetzen und die technischen Voraussetzungen hierzu zu schaffen. Andererseits galt es auch sicherzustellen, dass die Labormuster – im Besonderen der App-Client[10] – im Rahmen der Bereitstellung der Informationen nicht selbst unnötig personenbezogene Daten verarbeiten (privacy-by-design). Hierzu wurden Pseudo-Berechtigungen sowie Opt-In / Opt-Out Lösungen entwickelt, die es Nutzerinnen und Nutzern ermöglichen, nur diejenigen Daten dem App-Client zur Verfügung zu stellen, die Nutzerinnen und Nutzer zur Verfügung stellen möchten. Zudem wurden durch ein System zufälliger IDs und getrennter Prozesse sichergestellt, dass auch aus den an das PGuard Backend[11] übermittelten Daten ein möglicher Personenbezug auf ein technisch notwendiges Minimum reduziert wird.

2.2.6 Verwertungsoptionen

Verbraucherinnen und Verbraucher haben im Rahmen des Forschungsprojektes bestätigt, dass Interesse an einer Lösung zur Förderung des Selbstdatenschutzes besteht. Es konnte zudem im Rahmen von Umfragen und Gesprächen bestätigt werden, dass die in den Labormustern integrierten Funktionen den Bedarf der Verbraucherinnen und Verbraucher abbilden können.

Die den Verbraucherinnen und Verbrauchern auf dem aktuellen Markt zur Verfügung stehenden Angebote decken die im Rahmen des Forschungsprojekts entwickelten Ansätze nicht oder nur zum Teil ab.

[9] Siehe 4.8.2.
[10] Siehe 5.
[11] Siehe 4.11.1.2.

Insbesondere fehlt es an einer leicht verständlichen Aufbereitung der Datenverarbeitungen, die es Verbraucherinnen und Verbrauchern ermöglicht, eine selbstbestimmte, eigene Entscheidung zu treffen. Entsprechend bieten aktuelle Lösungen auch meist nur die Empfehlung, auf eine App zu verzichten. Die im Forschungsprojekt entwickelten, konkreten Handlungsempfehlungen, die soweit möglich eine Weiternutzung betroffener Apps ermöglichen soll, fehlen.

Grundsätzlich bestehen für die im Forschungsprojekt entwickelten Ergebnisse Verwertungsoptionen. Eine Verwertung ist in unterschiedlichen Zielmärkten vorstellbar. Auf Basis der Analyse der potentiellen Verwertungsoptionen wird jedoch davon ausgegangen, dass eine Verwertung in nur einem einzigen Zielmarkt nicht tragfähig sein könnte.

2.3 Besonderer Dank

Unser besonderer Dank gilt dem Bundesministerium für Bildung und Forschung für die Förderung des Vorhabens sowie dem Projektträger VDI/VDE Innovation + Technik für die Betreuung. Darüber hinaus möchten wir uns bei einer Reihe von Expertinnen und Experten bedanken, die im Laufe der Projektzeit wertvollen Input zum Projekt gegeben haben. Hierzu zählen insbesondere Kristin Benedikt (Bayerisches Landesamt für Datenschutzaufsicht), Constanze Gaßmann (Verbraucherzentrale Brandenburg), Florian Glatzner (Verbraucherzentrale Bundesverband), Dr. Carsten Hayungs (Bundesministerium der Justiz und für Verbraucherschutz), Matthias Kammer (DIVSI), Thomas Kranig (Bayerisches Landesamt für Datenschutzaufsicht), Erik Krempel (Fraunhofer), Dr. Dirk Lorenz (Stiftung Warentest), Petra Maid (DATEV eG), Frederik Richter (Stiftung Datenschutz), Miriam Ruhenstroth (mobilsicher), Matthias Spielkamp (mobilsicher), Dr. Claus-Dieter Ulmer (Deutsche Telekom), Simone Vintz (Stiftung Warentest), sowie den Mitgliedern des Arbeitskreis Datenschutz des Bitkom und den Teilnehmerinnen und Teilnehmer des Interdisziplinären Workshops „Privacy, Datenschutz & Surveillance“ des Humboldt Institut für Internet und Gesellschaft.

3 Forschungshintergrund und Gegenstand

Das Ziel dieses Forschungsvorhabens bestand darin, die wissenschaftlichen Grundlagen für eine PrivacyGuard App (kurz: PGuard App) und ein Webportal zu entwickeln und diese Anwendungen im Rahmen eines Labormusters im Hinblick auf deren Praktikabilität zu validieren.

Mithilfe der PGuard App (bzw. „DATENSCHUTZscanner“) und des Webportals sollten Nutzerinnen und Nutzer erstens Möglichkeiten gegeben werden, ohne großen Aufwand eine konkrete Risikobewertung der auf ihrem mobilen Endgerät installierten Apps vorzunehmen und hierbei die Vertrauenswürdigkeit und den konkreten Datenumgang installierter Apps leicht verständlich zu beurteilen. Zweitens sollten die App und das Webportal die Nutzerinnen und Nutzer darin unterstützen, auf der Grundlage der Risikobewertung zu handeln. Handlungen sollten sich etwa in einer Deinstallation einer App mit einem hohen Risiko, in einer Optimierung der Datenschutzgrundeinstellungen auf dem Smartphone, in dem Ausstieg aus dem Tracking eines Drittanbieters oder in der Deaktivierung bestimmter Funktionen der jeweiligen Apps widerspiegeln.

Vor dem Hintergrund, dass die Downloadzahlen von Apps in den kommenden Jahren weiter ansteigen werden und dass die Hindernisse für einen wirksamen Selbstdatenschutz auch durch geänderte rechtliche Rahmenbedingungen nicht vollends beseitigt sein werden, wurde ein fünffacher Handlungsbedarf ermittelt und zum Gegenstand dieses Forschungsvorhabens:

1. Das Bewusstsein der Verbraucherinnen und Verbraucher für Datenschutz- und Datensicherheitsgefahren bei der Verwendung von Apps muss gesteigert werden.
2. Verbraucherinnen und Verbrauchern muss eine einfache und praktikable Lösung zur Verfügung gestellt werden, mit deren Hilfe sie die konkreten Risiken der Apps, die auf ihrem Smartphone installiert sind, einfach und verständlich beurteilen können.
3. Verbraucherinnen und Verbraucher sollten konkrete Handlungsanleitungen erhalten, wie sie auf der Grundlage der Risikobewertungen handeln können, um hierdurch einen wirksamen Selbstdatenschutz praktizieren zu können.
4. Es müssen Möglichkeiten geschaffen werden, damit die Unternehmen, die Datenschutz und Datensicherheit ernst nehmen, auch vom Markt honoriert werden.
5. Es muss Druck auf solche App-Anbieter ausgeübt werden, die Datenschutz und Datensicherheitsstandards unterlaufen (die „schwarzen Schafe“).

Damit die App und das Webportal diese Zielsetzungen erreichen und überdies alltagstauglich und durch Laien einsetzbar sind, wurden folgende Mindestanforderungen untersucht und erforscht:

- nachvollziehbare und belastbare Daten und Auswertungslogiken Datenschutzerklärungen (DSE), (de jure),
- nachvollziehbare und belastbare Daten und Auswertungslogiken bezüglich des tatsächlichen Verhaltens der Apps, (de facto),
- Gestaltung, Nutzerführung, Risikobewertung und Handlungsempfehlungen der PGuard App intuitiv verständlich umsetzen,

- ggf. die Nutzerinnen und Nutzer durch Anstöße (Nudges) zu einem möglichst risikobewussten Handeln animieren,
- Bereitstellung zielführender Informationen bei Vermeidung pauschalierter Gesamtbeurteilungen einer App; vielmehr sollte eine Bevormundung ausgeschlossen und jeder Einzelne in die Lage versetzt werden, aufgeklärt und selbstbestimmt auf Basis seiner persönlichen Präferenzen eine individuelle Beurteilung zu erhalten,
- zunehmend (semi-)automatisierte Analysen um die große und stets wachsende Zahl an verfügbaren Apps für Nutzerinnen und Nutzer adressieren zu können.

Die Gliederung dieses Abschlussbericht orientiert sich im Wesentlichen an den aufgelisteten Herausforderungen.

Kapitel 4 befasst sich zunächst mit den grundlegenden Fragestellungen des Forschungsprojekts; das umfasst Eckpfeiler einer intuitiven und verständlichen Informationsaufbereitung, ebenso wie grundsätzliche Forschung zur Generierung belastbarer und nachvollziehbarer semantischen und technischen Analyseergebnisse.

Kapitel 5 widmet sich sodann den Aspekten, die sich aus besonderen Begebenheiten der Entwicklung eines App-Clients ergeben; hierbei wird auch auf etwaige Herausforderungen eingegangen, die sich aus Besonderheiten des jeweiligen Betriebssystems ergeben können.

Kapitel 6 führt die Forschungsergebnisse hinsichtlich sonstiger Darstellungsoptionen aus; dies umfasst sowohl die Darstellung der Informationen über ein Webportal als auch über alternative Ansätze, wie zum Beispiel ein Browser-Plugin.

Kapitel 7 stellt mögliche Herausforderungen im Rahmen einer Verwertung der Forschungsergebnisse dar; hierbei werden sowohl unterschiedliche Märkte als auch unterschiedliche, denkbare Verwertungsoptionen erläutert und mit einer Einschätzung tatsächlicher Erfolgsaussichten versehen.

4 Generelle Forschung und Ergebnisse

4.1 Einleitung

Dieses Kapitel widmet sich all jenen Forschungsergebnissen, die für das gesamte Forschungsvorhaben abstrakt – unabhängig eines konkreten Einsatzes in einem der Labormuster – anzustellen waren. Dies umfasst zunächst die Erfassung der Bedarfe, Interessen und des Wissensstandes der Verbraucherinnen und Verbraucher.

Die Kenntnis der für Verbraucherinnen und Verbraucher relevanten Aspekte, reflektiert in den Informationstexten, ermöglichte es zudem die Forschung im Bereich der Analysen zu fokussieren. Die generelle Forschung im Bereich der technischen[12] und semantischen[13] Analyse werden somit ebenfalls in diesem Kapitel dargestellt.

Zudem stellt dieses Kapitel die für die Labormuster notwendige Hintergrundtechnik (PGuard Backend[14]) dar. Dieses zentrale PGuard Backend ermöglicht die Aufbereitung und Darstellung der durch die Analysen gewonnenen Erkenntnisse in unterschiedlichen Labormustern.

4.2 Übergeordnete Methodik: Online-Befragungen

Im Folgenden wird die Methodik der beiden Online-Befragungen mit Verbraucherinnen und Verbrauchern bzw. potentiellen Nutzerinnen und Nutzern der DATENSCHUTZscanner-Anwendungen beschrieben. Neben einer inhaltlichen Einführung werden Datenerhebung, Stichprobenziehung sowie Hinweise zur Darstellung der Ergebnisse dargelegt:

Im August 2016 und Juni 2018 wurden zwei Online-Befragungen durchgeführt. Ziel der ersten Befragung war es, eine Übersicht über die Kompetenzen und Interessen von Smartphone-Nutzerinnen und -Nutzern in Bezug auf das Thema „Schutz personenbezogener Daten“ im Smartphone zu gewinnen und ein besseres Bild über potentielle Nutzerinnen und Nutzer der PrivacyGuard-App zu erhalten. Der Fokus der zweiten Befragung lag auf einer weiteren Untersuchung der Interessensgebiete und Herausforderungen, die Verbraucherinnen und Verbraucher bei der Benutzung ihres Smartphones haben.

Beide Umfragen wurden als Online-Fragebogen an ein repräsentatives Smartphone-Nutzer-Sample der deutschen Bevölkerung versendet und konnten im Webbrowser, auf dem Smartphone oder Tablet ausgefüllt werden.

4.2.1 Datenerhebung und Stichprobenziehung

Die Datenerhebung wurde in beiden Fällen durch das Befragungsunternehmen OmniQuest GmbH durchgeführt. Die Datenauswertung und Berichterstellung erfolgten durch die Quadriga Hochschule Berlin. An beiden Studien nahmen jeweils 1.000 Teilnehmerinnen und Teilnehmer im Alter von 14 bis 86 Jahren teil. Die Befragungspopulationen entsprechen der repräsentativen Smartphone-Nutzer-Bevölkerung, wobei sich die Quoten in Bezug auf Alter, Geschlecht und Wohnort in der zweiten Befragung an

[12] Siehe 4.9.
[13] Siehe 4.10.
[14] Siehe 5.5.3.

den Werten der ersten Befragung orientieren. Eine Übersicht über die Stichprobenzusammensetzung nach den soziodemografischen Merkmalen sowie die Fragebögen finden sich im Appendix (Kapitel 9).

4.2.2 Hinweise zur Darstellung der Ergebnisse

Die in den folgenden Kapiteln dargestellten Ergebnisse der Befragungen sind zum Teil auf ganze Zahlen gerundet, sodass es vorkommen kann, dass sich die Anteilswerte nicht immer zu 100% aufsummieren. Bei Fragen mit einer Möglichkeit zur Mehrfachnennung können die aufaddierten Nennungen überdies 100% überschreiten.

Des Weiteren ist zu beachten, dass aufgrund der Befragungslogik nicht allen Befragten der komplette Fragebogen vorgelegt wurde, sodass die Größe der befragten Grundgesamtheit zwischen den einzelnen Fragen variieren kann. Ein Beispiel sind die „Informationstexte“[15], die in einer randomisierten Fragebatterie abgefragt wurden. Abweichungen können sich weiterhin durch verweigerte Antworten zu bestimmten Fragestellungen ergeben. Die jeweilige Stichprobengröße pro Frage kann den zugehörigen Abbildungen und Tabellen entnommen werden.

4.3 Allgemeiner Kenntnisstand von Verbraucherinnen und Verbrauchern

Im Rahmen der ersten Befragung wurden die Datenschutzkompetenzen von Verbraucherinnen und Verbrauchern untersucht. Hierbei wurden sowohl die selbsteingeschätzte Datenschutzkompetenz, als auch die tatsächliche Datenschutzkompetenz in einem Quiz abgefragt. Das Quiz beinhaltete dabei Fragen zu:

1. Einstellungen und Handlungsweisen beim Smartphone und
2. Verbraucherrechten.

Im Folgenden werden die Ergebnisse der Befragung vorgestellt.

4.3.1 Selbsteinschätzung Wissensstand

Die Mehrheit der Teilnehmerinnen und Teilnehmer (52%) schätzt den eigenen Wissensstand zum Thema Datenschutz im Smartphone als mittelmäßig ein. Weitere 32% geben an, nur über ein geringes oder sehr geringes Wissen zu verfügen, während nur 16% angeben, dass sie über ein hohes oder sehr hohes Wissen verfügen.

[15] Siehe 4.8.2.

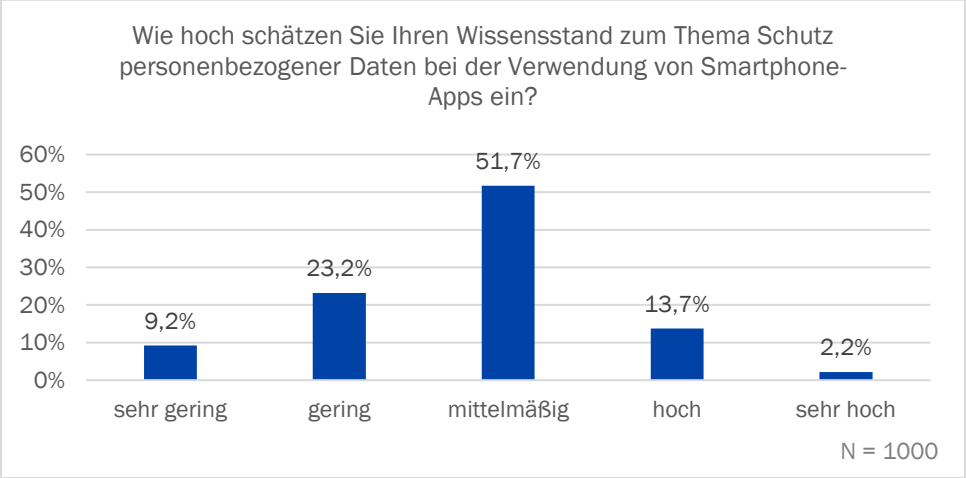


Abbildung 1: Selbsteinschätzung des Wissensstands

4.3.2 Tatsächlicher Wissenstand (Quiz)

Insgesamt wurden vier Fragen zum tatsächlichen Wissensstand der Verbraucherinnen und Verbraucher gestellt. Die erste Frage bezieht sich auf datenschutzfreundliche Einstellungen, die zweite auf Rechte und Pflichten beim Lesen der Datenschutzerklärungen, die dritte auf die Beschaffenheit von Einwilligungen und die vierte Frage auf die Sprache, in der die Datenschutzerklärung verfasst sein sollte. Im Folgenden werden die Fragen zuerst einzeln ausgewertet. Danach wird das Gesamtabschneiden der Teilnehmerinnen und Teilnehmer im Quiz beschrieben und der Zusammenhang zwischen Selbsteinschätzung und tatsächlichem Wissen hergestellt. Im letzten Schritt werden Unterschiede zwischen soziodemografischen Gruppen untersucht und dargestellt.

4.3.2.1 Quiz-Frage 1: Datenschutzfreundliche Einstellungen

Die unten abgebildete Grafik zeigt die erste Quiz-Frage zu datenschutzfreundlichen Einstellungen nebst korrekter Antwort (grün) und den drei inkorrekten Antworten (rot). Nur 36% der Befragten können die Frage nach datenschutzfreundlichen Einstellungen im Smartphone richtig beantworten und wissen, dass durch das Deaktivieren von Ortungsdiensten und Schnittstellen der Datenschutz verbessert werden kann. 46% sind fälschlicherweise der Ansicht, dass bereits ein aktueller Virenscanner garantiert, dass Daten nicht ungewollt verschickt werden. Weitere 11% vermuten, dass ein aktuelles Betriebssystem vor ungewollter Datenverarbeitung schützt und 7% glauben, dass Voreinstellungen von Smartphones bereits datenschutzgerecht sind.

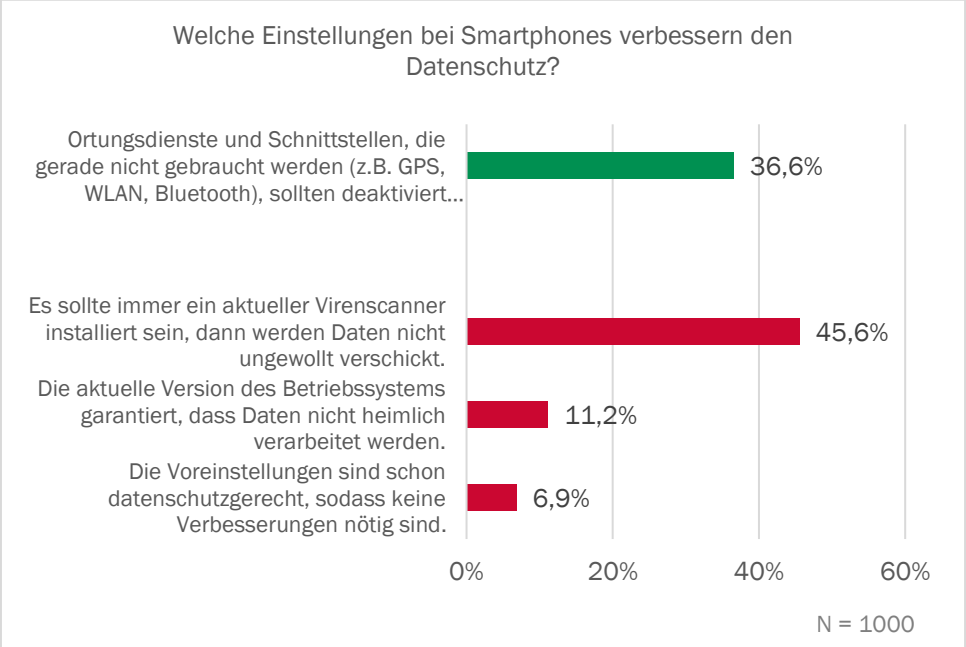


Abbildung 2: Quiz-Frage 1: Datenschutzfreundliche Einstellungen

4.3.2.2 Quiz-Frage 2: Lesen der Datenschutzerklärung

Die zweite Quiz-Frage beschäftigt sich mit dem Lesen der Datenschutzerklärung von neuinstallierten Apps und die unten abgebildete Grafik gibt die Fragestellung nebst korrekter Antwort (grün) und den drei inkorrekten Antworten (rot) wider. Die Mehrheit der Befragten (58%) weiß, dass das Lesen des „Kleingedruckten“ in den Datenschutzerklärungen nicht verpflichtend ist, ist sich jedoch bewusst, dass das Lesen der Datenschutzerklärung ratsam ist. 29% der Teilnehmerinnen und Teilnehmer vermuten fälschlicherweise, dass das Lesen verpflichtend ist und man sich strafbar macht, wenn man Datenschutzerklärungen mit „gelesen“ bestätigt, ohne dies tatsächlich getan zu haben. 13% sind zwar darüber informiert, dass das Lesen nicht verpflichtend ist, jedoch vermutet hiervon jeweils die Hälfte, dass dies entweder nur auf kostenlose Angebote zutrifft oder eine Unterschrift die Zustimmung erst bindend macht.

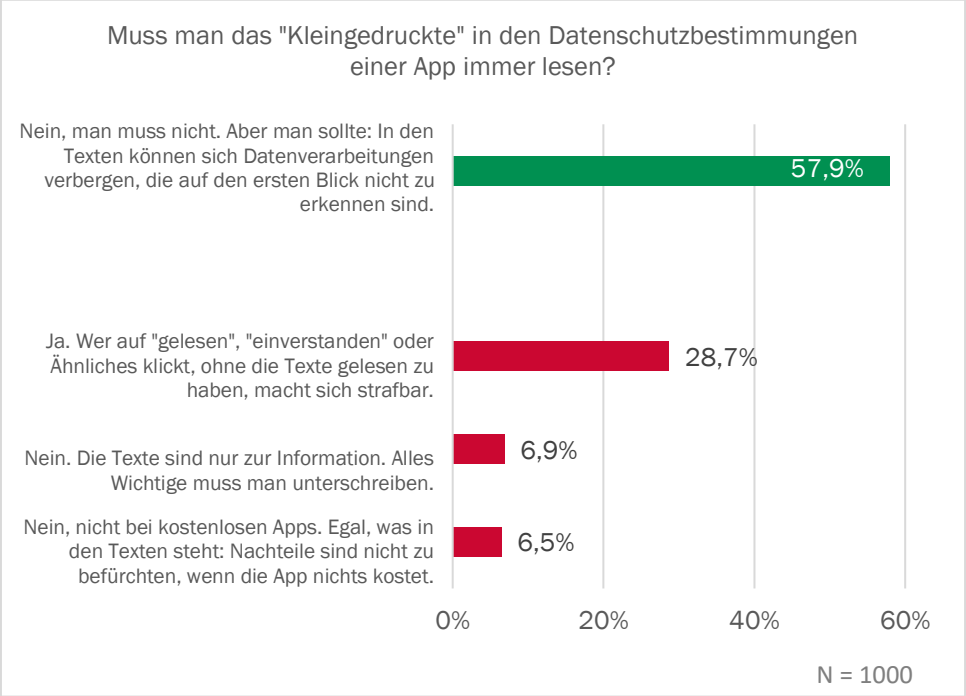


Abbildung 3: Quiz-Frage 2: Lesen der Datenschutzerklärung

4.3.2.3 Quiz-Frage 3: Einwilligung zur Datenverarbeitung

Die unten abgebildete Grafik zeigt die dritte Quiz-Frage zur Einwilligung zur Datenverarbeitung nebst korrekter Antwort (grün) und den drei inkorrekten Antworten (rot). Nur 32% der Teilnehmerinnen und Teilnehmer wissen, dass die Einwilligung zu einer Datenverarbeitung freiwillig und im Nachhinein widerrufbar ist. 33% der Befragten gingen fälschlicherweise davon aus, dass einer Datenschutzerklärung explizit zugestimmt werden muss. Ebenso glauben diese 33% die Einwilligung gelte für alle Datenverarbeitungen und unabhängig von der Zweckmäßigkeit. 19% vermuten, dass die Einwilligung mit der ersten Nutzung beginnt und somit unabhängig davon ist, ob sie ausdrücklich erteilt wurde oder nicht. Weiteren 16% ist zwar bewusst, dass einer Datenschutzerklärung zugestimmt werden muss, jedoch vermuten sie, dass ein Widerruf im Nachhinein nicht mehr möglich ist.

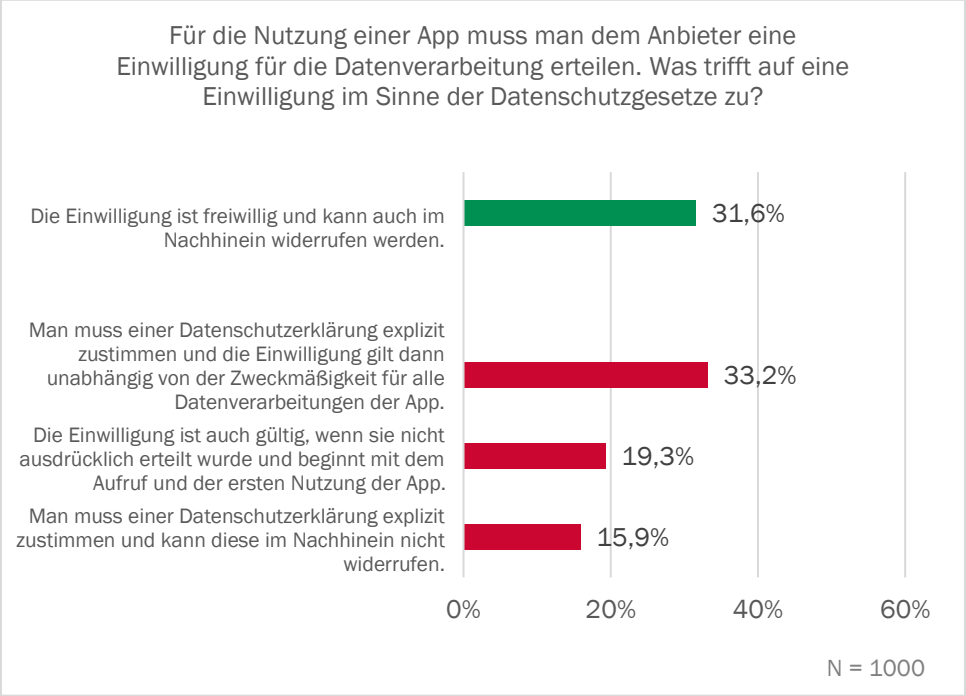


Abbildung 4: Quiz-Frage 3: Einwilligung zur Datenschutzerklärung

4.3.2.4 Quiz-Frage 4: Sprache der Datenschutzerklärung

Die letzte Quiz-Frage beschäftigt sich mit der Sprache der Datenschutzerklärungen. So ist geregelt, dass die Datenschutzerklärung in verständlicher Sprache vorliegen muss; mithin die Datenschutzerklärung einer App, die sich an deutsche Nutzerinnen und Nutzer richtet, in deutscher Sprache vorliegen muss. Im Rahmen des Forschungsprojektes und für den Entwurf der vierten Quiz-Frage wurde sich für eine strenge Auslegung des Rechtstextes, insbesondere im Hinblick auf die mit Ende des Projektes anzuwendende Datenschutzgrundverordnung, entschieden. Die Datenschutzgrundverordnung hat die Transparenzpflichten flächendeckend verschärft. Die Anforderungen an Datenschutzerklärungen hinsichtlich Verständlichkeit und einfacher Sprache wurden konkretisiert.[16]

Die unten abgebildete Grafik zeigt die Fragestellung nebst korrekter Antwort (grün) und den drei inkorrekten Antworten (rot). Die Mehrheit der Befragten beantwortet diese Frage korrekt (53%). 27% vermuten fälschlicherweise, dass die Datenschutzerklärung in einer Amtssprache der Europäische Union verfasst sein muss. 12% geben an, dass es für die Sprache der Datenschutzerklärung keine Regeln gibt und 9% glauben, dass die Sprache des App-Anbieters über die Sprache der Datenschutzerklärung entscheidet.

[16] Eine Datenschutzerklärung, die sich an einen Markt mit einer von den dort befindlichen Betroffenen primär gesprochenen und verstandenen Sprache oder mit dezidierte Amtssprache(n) wendet (in Deutschland wäre dies Deutsch), erscheint nur verständlich im Sinne der rechtlichen Vorgaben, wenn diese eben diese primär gesprochene und verstandene Sprache oder die dezidierten Amtssprache(n) abbildet.

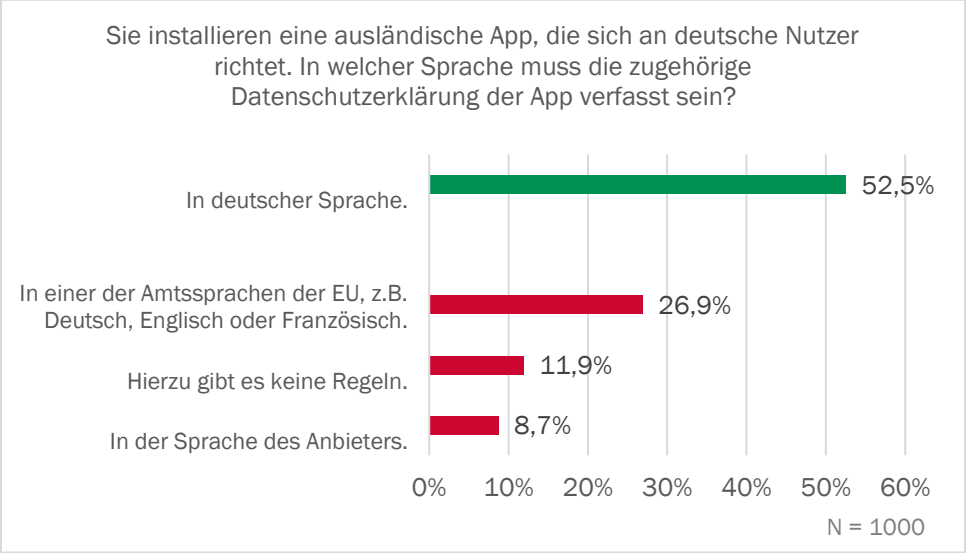


Abbildung 5: Quiz-Frage 4: Sprache der Datenschutzerklärung

4.3.2.5 Gesamtanzahl korrekter Antworten im Quiz

Im Schnitt können Teilnehmerinnen und Teilnehmer etwas weniger als die Hälfte aller Quizfragen korrekt beantworten. Diese Beobachtung deckt sich mit den Ergebnisse von Trepte und Masur (2015), die die Privatheitskompetenz ihrer Befragten ebenfalls mit weniger als der Hälfte der korrekten Antworten messen, jedoch unterscheidet sich die Art deren Fragebogens stark von dem hier behandelten Fragebogen, der seinen Fokus auf Datenschutz und Apps richtet.[17]

Die Verteilung der Quiz-Performance über alle Teilnehmerinnen und Teilnehmer und die vier Quiz-Fragen hinweg stellt sich wie folgt dar: 10% der Teilnehmerinnen und Teilnehmer können keine einzige Frage korrekt beantworten. 29% der Befragten können eine einzige Frage korrekt beantworten. 37% der Teilnehmerinnen und Teilnehmer können zwei von vier Quiz-Fragen korrekt beantworten, 19% drei von vier und immerhin 4% beantworten alle vier Fragen korrekt.

[17] Trepte, S. & Masur, P. K. (2015). Privatheit im Wandel. Eine repräsentative Umfrage zur Wahrnehmung und Beurteilung von Privatheit (Bericht vom 18. Juni 2015). Stuttgart: Universität Hohenheim. Abgerufen von: https://www.uni- hohenheim.de/fileadmin/einrichtungen/psych/Team_MP/Berichte/Bericht_- _Privatheit_im_Wandel_2014-06-18.pdf

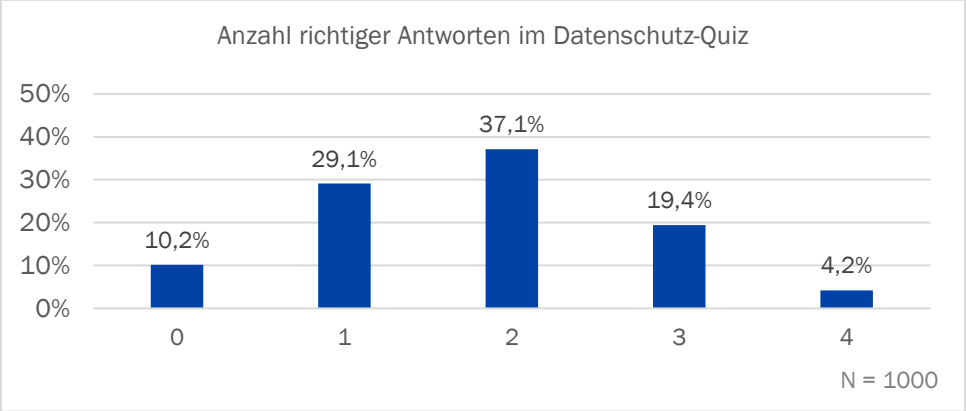


Abbildung 6: Anzahl richtig beantworteter Quiz-Fragen

4.3.3 Zusammenhang zwischen Selbsteinschätzung und tatsächlichem Wissensstand

Zwischen der Selbsteinschätzung des eigenen Datenschutzwissensstands und der tatsächlichen Performance im Quiz gibt es keinen Zusammenhang. Die Korrelation der beiden Variablen ist nahe Null und insignifikant.

Die nachstehende Abbildung 7 zeigt einen Bubble-Chart, der den Anteil der Teilnehmer in den jeweiligen Kategorien der Selbsteinschätzung und der Anzahl der richtig beantworteten Fragen anzeigt.[18] Insgesamt schätzen 29% der Teilnehmerinnen und Teilnehmer ihren Wissensstand realistisch[19] ein. 36% schätzen ihr Wissen zu niedrig ein und schneiden im Quiz besser ab, als sie sich zutrauen und die restlichen 35% überschätzen sich selbst und schneiden im Quiz schlechter ab, als sie vermuten.

[18] Auf der horizontalen Achse wird der selbsteingeschätzte Wissenstand abgetragen und auf der vertikalen Achse die Anzahl der korrekten Antworten im Quiz (tatsächlicher Wissensstand). Die Kreise zeigen anhand ihrer Lage auf dem Gitter jeweils an, welcher selbsteingeschätzte und welcher tatsächliche Wissensstand zutreffen. Zusätzlich zeigt die Größe der Kreise den prozentualen Anteil der Befragten an, auf den die Kombination von selbsteingeschätztem und tatsächlichen Wissensstand zutrifft.

[19] Die Definition von realistisch lässt sich an folgenden Beispielen erklären: Ein Teilnehmer schätzt seinen Wissensstand mit „mittelmäßig“ ein und beantwortet genau die Hälfte der Fragen. Oder ein Teilnehmer schätzt seinen Wissensstand mit „sehr gering“ ein und beantwortet keine Quiz-Frage korrekt.

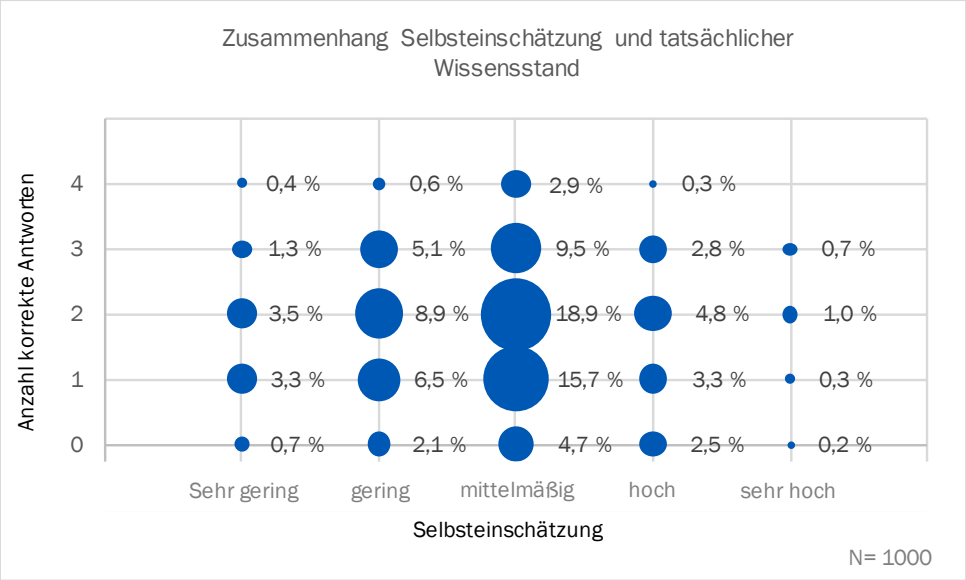


Abbildung 7: Zusammenhang der Selbsteinschätzung mit dem tatsächlichen Wissensstand

4.3.4 Wissensstand und Unterschiede zwischen Bevölkerungsgruppen

Zusätzlich zu den übergeordneten Ergebnissen zum Wissensstand wurden Unterschiede im Abschneiden der Teilnehmerinnen und Teilnehmer aufgrund ihres Geschlechts, Alters, Bildungsstands und Einkommens untersucht. Die folgenden Abschnitte fassen die Ergebnisse zusammen.

4.3.4.1 Geschlecht

Die durchschnittliche Selbsteinschätzung zwischen Männern und Frauen unterscheidet sich signifikant. Frauen haben eine durchschnittliche Selbsteinschätzung von 2,9 (Skala von 1=„sehr gering“ bis 5=„sehr hoch“), wohingegen Männer ihren Wissensstand im Durchschnitt nur mit 2,6 einschätzen.

Bei dem tatsächlichen Abschneiden in den Quiz-Fragen gibt es jedoch keinen signifikanten Unterschied. So ist die Anzahl der korrekten Antworten bei Frauen und Männern im Durchschnitt 1,8 bzw. 1,76, das heißt sie beantworten im Schnitt keine zwei Fragen korrekt.

Dies steht im Gegensatz zu einer Befragung von Trepte und Masur (2015), die feststellen, dass Männer über eine höhere Privatheits- und Datenschutzkompetenz verfügen als Frauen.[20] Jedoch werden in deren Studie andere Kompetenzen als in dem hier behandelten Datenschutz-Quiz erhoben, das sich vor allem auf die Nutzung von Smartphones und damit verbundene datenschutzrechtliche Aspekte fokussiert.

4.3.4.2 Alter

Der Zusammenhang zwischen dem selbsteingeschätzten Wissensstand der Teilnehmerinnen und Teilnehmer und ihrem Alter ist zwar signifikant und negativ, jedoch gering (ρ=-0.17; $p<0,01$). Das heißt,

[20] Trepte, S. & Masur, P. K. (2015). Privatheit im Wandel. Eine repräsentative Umfrage zur Wahrnehmung und Beurteilung von Privatheit (Bericht vom 18. Juni 2015). Stuttgart: Universität Hohenheim. Abgerufen von: https://www.uni- hohenheim.de/fileadmin/einrichtungen/psych/Team_MP/Berichte/Bericht_- _Privatheit_im_Wandel_2014-06-18.pdf

dass mit steigendem Alter der selbsteingeschätzte Wissensstand gering fällt. Zwischen der Quiz-Performance und dem Alter gibt es keinen Zusammenhang.

4.3.4.3 Bildung

Sowohl zwischen dem selbsteingeschätzten Wissensstand als auch zwischen der tatsächlichen Quiz-Performance besteht kein Zusammenhang mit dem Bildungsstand der Befragten.

4.3.4.4 Einkommen

Zwischen dem selbsteingeschätzten Wissensstand und dem Einkommen besteht ein geringer, positiver und signifikanter Zusammenhang (ρ=-0.17; p<0,01), das heißt, dass mit wachsendem Einkommensniveau auch das selbsteingeschätzte Wissen geringfügig höher ist. Zwischen tatsächlicher Performance und Einkommen kann jedoch keine Korrelation beobachtet werden.

4.3.5 Zusammenfassung

Zusammenfassend lässt sich aus den Ergebnissen der Befragung ableiten, dass beim Thema Datenschutz bei Smartphone-Apps Nachholbedarf bezüglich der Kompetenz von Verbraucherinnen und Verbrauchern existiert. Im Durchschnitt schätzen Verbraucherinnen und Verbraucher ihre Datenschutz-Kompetenz mittelmäßig ein und diese mittelmäßige Kompetenz wird durch das Abschneiden im Wissensquiz auch bestätigt.

4.4 Motivation von Verbraucherinnen und Verbrauchern Selbstdatenschutzmaßnahmen zu ergreifen

In der ersten Befragung wurden außerdem Maßnahmen untersucht, die Verbraucherinnen und Verbraucher ergreifen, um ihre Daten zu schützen.

4.4.1 Übersicht über die Datenschutzmaßnahmen und Ergebnisse

Die untersuchten Maßnahmen umfassten sowohl einen Smartphone-spezifischen Teil, wie das Deinstallieren von Apps, das Entziehen von App-Berechtigungen oder das Lesen einer Datenschutzerklärung einer App, als auch Maßnahmen, die im Allgemeinen auf datenschutzfreundliches Verhalten im Internet abzielen. Zu den letzteren zählen zum Beispiel die Nutzung von Betroffenenrechten, die Verwendung von Pseudonymen, das Entfernen von Verlinkungen, zum Beispiel in sozialen Netzwerken, oder das Verändern von Gerätekennungen. Außerdem wurde abgefragt, ob eine Datenschutzerklärung auf einer Webseite gelesen wurde.

24% der Befragten geben an, bisher noch keine Datenschutzmaßnahme ergriffen zu haben. Somit haben 76% der Teilnehmerinnen und Teilnehmer bereits etwas für ihren Datenschutz unternommen (mit einem Durchschnittswert von 1,8 Maßnahmen pro Person). Die Mehrheit der Befragten (32%) gibt an, eine Maßnahme ergriffen zu haben. 44% ergreifen sogar mehr als eine der genannten Maßnahmen.

Die folgende Abbildung 8 listet die einzelnen Datenschutzmaßnahmen auf und gibt die Häufigkeit der Teilnehmerinnen und Teilnehmer an, die diese Maßnahme in der Vergangenheit genutzt haben, um ihren Datenschutz zu verbessern.

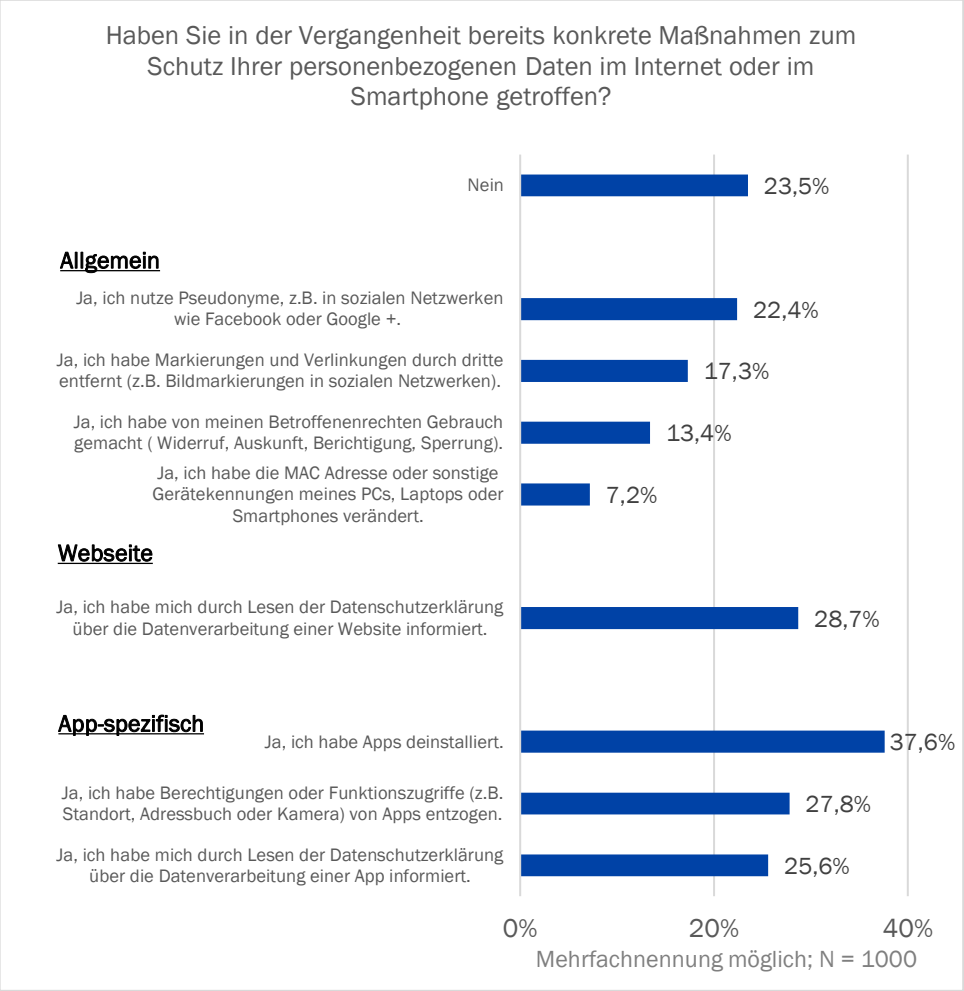


Abbildung 8: Maßnahmen zum Selbstdatenschutz

Die am häufigsten ergriffene Maßnahme zum Schutz personenbezogener Daten ist mit 38% die Deinstallation von Apps. Auch das Lesen einer Webseiten-Datenschutzerklärung wird von 29% der Teilnehmerinnen und Teilnehmer als Maßnahme genannt, wobei der Wert für App-Datenschutzerklärungen mit 26% geringer ist. Auch Berechtigungen, wie Standort-, Adressbuch- oder Fotozugriffe, werden häufig entzogen (28%).

Allgemeine Maßnahmen, die unabhängig vom Smartphone auch für Internetdienste zutreffen, werden ebenfalls ergriffen. So nutzen 22% der Teilnehmerinnen und Teilnehmer Pseudonyme und 17% haben Markierungen und Verlinkungen entfernt. Betroffenenrechte werden nur von 13% der Teilnehmerinnen und Teilnehmer genutzt und die Veränderung von Gerätekennungen ist mit 7% die am wenigsten genutzte Maßnahme.

Zu den aufgeführten Ergebnissen gibt es sehr wenige vergleichbare Studienergebnisse. Eine Studie von TNS Emnid und dem vzbv stellte 2015 fest, dass 15% der Befragten die Datenschutzerklärungen eines

Dienstes „nie“ lesen. Die restlichen 84% der Befragten gaben zumindest an, dass sie die Datenschutzerklärungen „selten“ oder öfter lesen.[21] Ähnlich sind die Ergebnisse einer Befragung des Bitkom (2015). Hier gaben 67% der Teilnehmerinnen und Teilnehmer an, dass sie zumindest gelegentlich eine Datenschutzerklärung lesen.[22] Dieser Wert ist in unserer Befragung geringer.[23]

Eine weitere Studie des Bitkom (2015) fand heraus, dass 52% der Smartphone-Nutzerinnen und -Nutzer bereits Standortberechtigungen entzogen hatten. Dieser Wert ist höher als in unserer Befragung, da er auch den globalen Standortzugriff und nicht nur den App-spezifischen Standortzugriff einschließt. Bei Fotozugriffen (38%), Adressbuchzugriffen (25%) und Mikrofonzugriffen (17%) sind die Werte vergleichbar mit unseren Ergebnissen.[24]

4.4.2 Zusammenhang zwischen Maßnahmen zum Selbstdatenschutz und Wissensstand

Zwischen der Anzahl der Datenschutzmaßnahmen und der Performance im Datenschutzquiz besteht ein schwacher positiver Zusammenhang. Die Korrelation beträgt ρ=0,18 ($p<0.001$). Das heißt, je höher der tatsächliche Wissensstand der Teilnehmerinnen und Teilnehmer ist, desto mehr Datenschutzmaßnahmen geben sie an durchzuführen.

Außerdem beobachtet man zwischen der Anzahl der Datenschutzmaßnahmen und dem selbstgeschätzten Wissen einen mittleren, positiven Zusammenhang. Der Korrelationskoeffizient ist ρ=0,35 und statistisch signifikant ($p<0,001$). Das heißt je höher der selbsteingeschätzte Wissensstand, desto höher auch die Anzahl der Datenschutzmaßnahmen.

4.4.3 Maßnahmen zum Selbstdatenschutz und Unterschiede zwischen den Bevölkerungsgruppen

4.4.3.1 Geschlecht

Zwischen Frauen und Männern ist ein signifikanter Unterschied zu beobachten. So geben Frauen an, im Durchschnitt 2,03 Datenschutzmaßnahmen zu ergreifen. Bei Männern ist dieser Wert geringer und liegt bei 1,59. Dies spiegelt sich auch in den Ergebnissen einer Befragung des vzbv wider. In der Dimension „Lesen der Datenschutzerklärung“ ist der Anteil der Befragten, die Datenschutzerklärungen selten oder nie lesen, bei Männern höher als bei Frauen.[25]

[21] TNS Emnid & vzbv (2015). Datenschutz - Die Sicht der Verbraucherinnen und Verbraucher in Deutschland. Abgerufen von: http://www.vzbv.de/sites/default/files/downloads/Datenschutz_Umfrage-Sicht-Verbraucher-Ergebnisbericht-TNS-Emnid-Oktober-2015.pdf

[22] Bitkom (2015). Datenschutz in der digitalen Welt. Seite 10. Abgerufen von: https://www.bitkom.org/Presse/Anhaenge-an-PIs/2015/09-September/Bitkom-Charts-PK-Datenschutz-22092015-final.pdf

[23] Im Experiment von Elshout et al. (2016) und in der Feldstudie von Kettner et al. (2018) liegen die gemessenen Werte zum Leseverhalten von Datenschutzerklärungen sogar noch niedriger. Weitere Informationen zur Messung und der verhaltenswissenschaftlichen Literatur zur Einwilligung von Verbraucherinnen und Verbrauchern finden sich in Kettner, S.E., Thorun, C. & Vetter, M. (2018). Wege zur besseren Informiertheit: Verhaltenswissenschaftliche Ergebnisse zur Wirksamkeit des One-Pager-Ansatzes und weiterer Lösungsansätze im Datenschutz.

[24] Bitkom (2015). Datenschutz bei Smartphone-Apps im Blick behalten. Abgerufen von: https://www.bitkom.org/Presse/Presseinformation/Datenschutz-bei-Smartphone-Apps-im-Blick-behalten.html

[25] TNS Emnid & vzbv (2015). Datenschutz - Die Sicht der Verbraucherinnen und Verbraucher in Deutschland. Abgerufen von: http://www.vzbv.de/sites/default/files/downloads/Datenschutz_Umfrage-Sicht-Verbraucher-Ergebnisbericht-TNS-Emnid-Oktober-2015.pdf

4.4.3.2 Alter

Zwischen der Anzahl der Datenschutzmaßnahmen und dem Alter ist ebenfalls eine Tendenz zu erkennen. Es besteht eine schwach negative, signifikante Korrelation zwischen Datenschutzmaßnahmen und Alter in Jahren (ρ=-0,15, p<0,001), das heißt je höher das Alter, desto geringer die Anzahl der Maßnahmen. Befragte unter 30 Jahren ergreifen im Schnitt 2,08 Maßnahmen. Befragte über 60 Jahren nur 1,3 Maßnahmen.

Ähnliche Ergebnisse finden sich in der Befragung des Bitkom aus dem Jahr 2015. Hier wurde ein starker Alterseffekt bei den Privatsphäre-Einstellungen von App-Zugriffen beobachtet. So gaben 44% der unter 30-Jährigen an, bereits einen Zugriff auf ihre Fotos verweigert zu haben, der Anteil der über 65-Jährigen lag nur bei 4%.[26]

4.4.3.3 Bildung

Auch zwischen Bildungsniveau und Datenschutzmaßnahmen besteht ein Zusammenhang. So steigt die Anzahl der Maßnahmen mit dem höchsten Bildungsabschluss und man kann eine schwache, signifikante Korrelation beobachten (ρ=0.15, p<0,001). Befragte, die als ihren höchsten Bildungsabschluss mindestens das Abitur angeben, führen im Schnitt 2,15 Datenschutzmaßnahmen durch. Teilnehmerinnen und Teilnehmer mit einem niedrigeren Bildungsabschluss geben als Durchschnittswert 1,64 Maßnahmen an.

4.4.3.4 Einkommen

Die Anzahl der Datenschutzmaßnahmen und das monatliche Nettoeinkommen sind ebenfalls schwach positiv korreliert. So findet man einen Korrelationskoeffizienten von ρ=0,12 (p<0,001), das heißt je höher das Einkommen, desto höher die Anzahl der Datenschutzmaßnahmen. Dies stellt jedoch nur eine Tendenz dar.

4.4.4 Zusammenfassung

Zusammenfassend lässt sich basierend auf den Befragungsergebnissen feststellen, dass Verbraucherinnen und Verbraucher grundsätzlich motiviert sind, etwas für ihren Selbstdatenschutz zu unternehmen. Gerade Maßnahmen, die einfach und ohne technische Kompetenz durchführbar sind, werden von Verbraucherinnen und Verbrauchern ergriffen.

4.5 Bedarf von Verbraucherinnen und Verbrauchern an Tools zum Schutz ihrer personenbezogenen Daten

Sodann wurde der Bedarf von Verbraucherinnen und Verbrauchern an Tools zum Schutz ihrer personenbezogenen Daten erfragt. Zum einen wurde untersucht, welche Funktionen für Verbraucherinnen und Verbraucher bei einer Datenschutzanwendung einen Mehrwert bieten und zum anderen wurde überprüft, ob Verbraucherinnen und Verbraucher bereit wären, solche Tools zu installieren und zu nutzen. Hieraus leitet sich die (hypothetische) Nachfrage für den DATENSCHUTZscanner ab.

[26] Bitkom (2015). Datenschutz bei Smartphone-Apps im Blick behalten. Abgerufen von: https://www.bitkom.org/Presse/Presseinformation/Datenschutz-bei-Smartphone-Apps-im-Blick-behalten.html

4.5.1 Relevante Funktionen

Im Rahmen der ersten Befragung wurden Verbraucherinnen und Verbraucher gefragt, wie sie bestimmte Funktionen eines möglichen Datenschutz-Tools für das Smartphone bewerten.

Die nachstehende Abbildung 9 listet die Kernfunktionen eines möglichen Datenschutz-Tools auf und den Anteil der Befragten, die die unterschiedlichen möglichen Funktionen als „wichtig“ oder „sehr wichtig“ bewerten. Im Durchschnitt wird die Relevanz der Funktionen zwischen 3,5 und 4,17 eingeschätzt. Der Mittelwert der Bewertungen über alle neun Kernfunktionen und alle Teilnehmerinnen und Teilnehmer liegt bei 3,91 mit einem Median von 4 (dies entspricht der Bewertungskategorie „wichtig“).[27]

Die Funktion „Automatisches Anzeigen besonders kritischer Aspekte“ wurde von den Teilnehmerinnen und Teilnehmern im Durchschnitt mit der höchsten Wichtigkeit versehen. Ein geringeres Interesse hatten die Befragten hingegen an dem „Anzeigen der Bewertungen von Apps durch andere Nutzerinnen und Nutzer oder Institutionen“. Diese Funktion ist den Teilnehmerinnen und Teilnehmern am wenigsten wichtig.

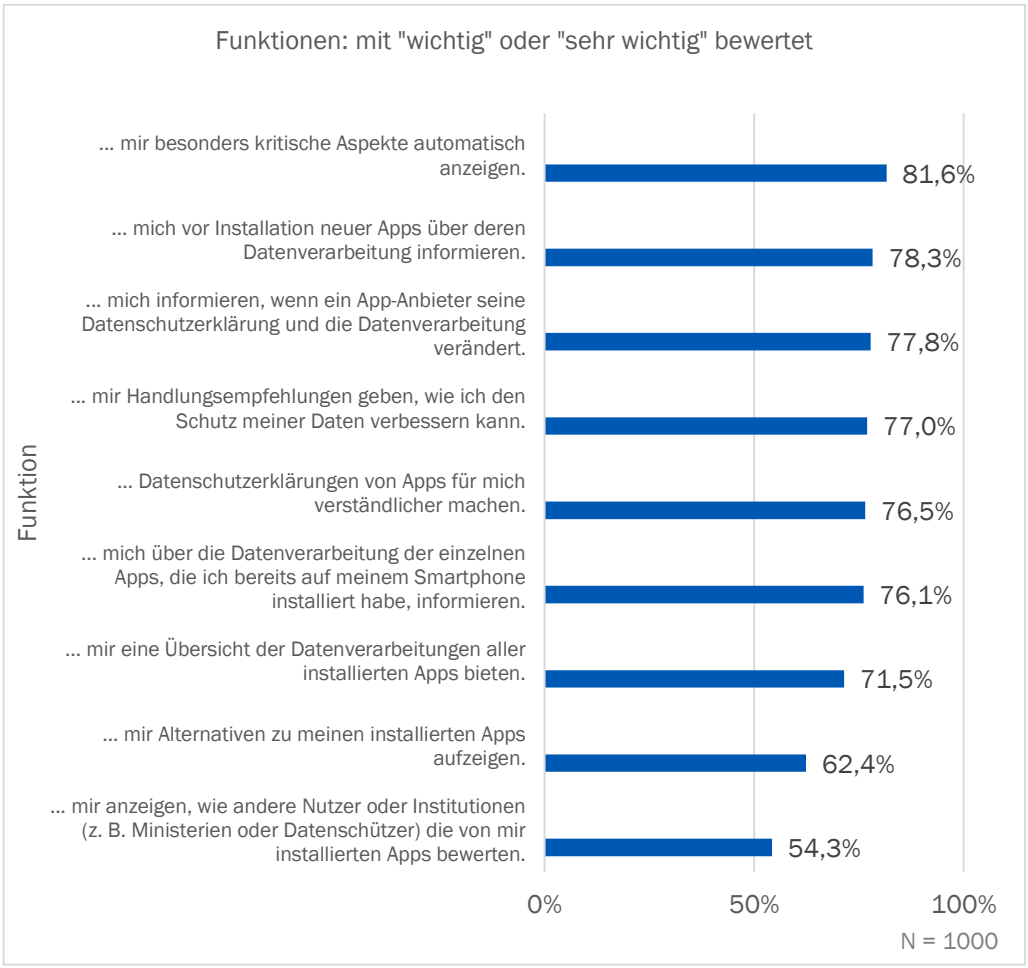


Abbildung 9: Interesse an einzelnen PrivacyGuard-Funktionen

[27] Die Funktionen wurden auf einer 5er-Skala von „sehr unwichtig“ (1) bis „sehr wichtig“ (5) bewertet.

4.5.2 Funktionen der Anwendung und Unterschiede zwischen den Bevölkerungsgruppen

Zusätzlich zu den Ergebnissen der Gesamtpopulation wurden diese Funktionen auch auf Unterschiede im Hinblick auf Alter, Geschlecht, Bildungsstand sowie Einkommen untersucht.

4.5.2.1 Geschlecht

Zwischen den einzelnen Funktionen ist kein systematischer Geschlechterunterschied zu beobachten. So bewerten Frauen und Männer die Funktionen größtenteils gleich (im Durchschnitt über alle Funktionen Frauen: 3,87 und Männer: 3,95). Das „Automatische Anzeigen besonders kritischer Aspekte“ wird von Männern im Durchschnitt höher bewertet als von Frauen. Dies trifft ebenfalls auf die Funktion „vor Installation neuer Apps über deren Datenverarbeitung informieren“ zu. Auch bei der Funktion „Datenschutzerklärungen verständlicher machen“ geben Männer im Durchschnitt eine signifikant höhere Bewertung ab.

4.5.2.2 Alter

Zwischen dem Alter und der durchschnittlichen Bewertung aller neun Funktionen ist eine geringe, positive Korrelation zu beobachten (ρ=0.15, p>0.001). Das heißt, je höher das Alter desto wichtiger werden die Funktionen bewertet. Einzig die Funktion „Anzeigen der Bewertungen von Apps durch andere Nutzer und Institutionen“ weist keinen Zusammenhang mit dem Alter auf.

4.5.2.3 Bildung und Einkommen

Zwischen Bildungsstand und Bewertung der App-Funktionen kann kein systematischer Unterschied festgestellt werden. Auch zwischen den Einkommensklassen und der Funktionenbewertung besteht kein signifikanter Zusammenhang.

4.5.3 Nachfrage und Zahlungsbereitschaft

Nach Auflistung und Bewertung der Kernfunktionen wurden die Teilnehmerinnen und Teilnehmer gefragt, ob sie eine solche Selbstdatenschutz-App, die die genannten Funktionen bietet, installieren würden. 61% der Befragten antworteten mit „ja“, 8% antworteten mit „nein“ und 31% waren sich nicht sicher, ob sie die App installieren würden.

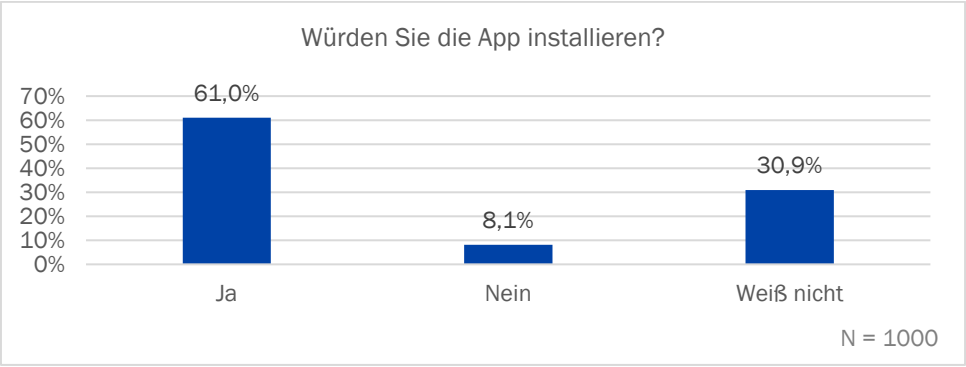


Abbildung 10: Nachfrage nach Datenschutz-App

Im Anschluss an die Frage zum Interesse an der Datenschutz-App mussten die Befragten angeben, wie viel sie für eine solche App bezahlen würden. Von den Teilnehmern, die angaben, dass sie eine Selbstdatenschutz-App installieren würden, gaben 68% einen Betrag über 0€ als Zahlungsbereitschaft an. Die Zahlungsbereitschaften spannen sich von 0€ bis 400€, mit einem Mittelwert von 5,85€ und einem Median von 3€. Die Gruppe der Befragten, die sich nicht sicher ist, ob sie eine Selbstdatenschutz-App installieren würde, hat eine geringere Zahlungsbereitschaft und nur 37% gaben an, mehr als 0€ für eine solche App zahlen zu wollen. Die Zahlungsbereitschaften liegen hier zwischen 0€ und 34€ mit einem Mittelwert von 1,97€ und einem Median von 0€.

Da die geplante Selbstdatenschutz-App mit einer Vielzahl spezieller Funktionen ausgestattet werden soll, gibt es in der Literatur wenig vergleichbare Ergebnisse zur Nachfrage und Zahlungsbereitschaft von Datenschutz-Apps. Eine Studie von TNS Emnid und vzbv (2015) fand jedoch heraus, dass 51% der Internetnutzerinnen und -nutzer bereit wären, für höchsten Datenschutz und Werbefreiheit bei Internetdiensten Geld auszugeben. 45% zeigten keine Zahlungsbereitschaft und 4% waren sich nicht sicher. Als explizite Zahlungsbereitschaft in Euro pro Monat gab die Hälfte der Befragten einen Wert bis zu 5 Euro an.[28]

4.5.4 Nachfrage, Zahlungsbereitschaft und Unterschiede zwischen den Bevölkerungsgruppen

Darüber hinaus wurden die Aspekte der Nachfrage und Zahlungsbereitschaft auf Unterschiede im Hinblick auf Geschlecht, Alter, Bildungsstand sowie Einkommen untersucht:

4.5.4.1 Geschlecht

Zwischen Frauen und Männern ist ein Unterschied in der Nachfrage nach der Selbstdatenschutz-App festzustellen. Die folgende Tabelle zeigt das Interesse an der App nach Geschlecht. Der Unterschied ist vornehmlich im höheren Anteil der unentschlossenen Männer zu beobachten.

Interesse an der PrivacyGuard-App

Geschlecht	Ja	Nein	Weiß nicht	Gesamt
Weiblich	63,84%	9,29%	26,87%	100%
Männlich	58,22%	6,93%	34,85%	100%

N=1000

Tabelle 1: Nachfrage nach Datenschutz-App nach Geschlecht

Die Zahlungsbereitschaft zwischen Männern und Frauen unterscheidet sich nicht. So sind Frauen zwar bereit, im Durchschnitt circa einen Euro mehr auszugeben (5,03 versus 4,09), jedoch ist der Unterschied nicht signifikant. In einer Umfrage von TNS Emnid und vzbv (2015) wurde zwischen Männern

[28] TNS Emnid & vzbv (2015). Datenschutz - Die Sicht der Verbraucherinnen und Verbraucher in Deutschland. Seite 10-11. Abgerufen von: http://www.vzbv.de/sites/default/files/downloads/Datenschutz_Umfrage-Sicht-Verbraucher-Ergebnisbericht-TNS-Emnid-Oktober-2015.pdf

und Frauen ein Unterschied in der Zahlungsbereitschaft für einen datenschutzfreundlichen und werbefreien Internetdienst gefunden,[29] jedoch sind die Funktionen der hier untersuchten Selbstdatenschutz-App sehr spezifisch und somit können die Ergebnisse nur bedingt verglichen werden.

4.5.4.2 Alter

Zwischen den Altersklassen und der Nachfrage ist ein gewisser Unterschied festzustellen. Die folgende Tabelle zeigt das Installationsinteresse nach Altersklassen an.

Interesse an der PrivacyGuard-App

Alter	Ja	Nein	Weiß ich nicht	Gesamt
14-19	70,79%	4,49%	24,72%	100%
20-29	68,05%	13,61%	18,72%	100%
30-39	67,90%	7,41%	18,34%	100%
40-49	54,50%	6,50%	24,69%	100%
50-59	59,39%	4,85%	39,00%	100%
60-69	49,49%	7,07%	35,76%	100%
70-79	58,72%	12,84%	43,43%	100%
80+	28,57%	0%	71,43%	100%
				N=1000

Tabelle 2: Nachfrage nach Datenschutz-App nach Altersklasse

In den Altersklassen unter 40 Jahren geben über zwei Drittel der Befragten an, die Selbstdatenschutz-App installieren zu wollen, doch auch in den höheren Altersklassen liegt der Anteil noch bei über 50%. Der Zusammenhang zwischen Nachfrage und Alter ist jedoch nicht linear. In der Gruppe der über 80-Jährigen gibt es außerdem ein geringeres Interesse an der App im Vergleich zu den anderen Altersklassen. Jedoch sollte diesem Ergebnis kein allzu großer Stellenwert beigemessen werden, da diese Altersgruppe nur einen geringen Anteil am Gesamtsample einnimmt.

Zwischen der Zahlungsbereitschaft und dem Alter kann kein Zusammenhang festgestellt werden.

4.5.4.3 Bildung und Einkommen

Zwischen Bildungsgrad und Einkommen kann kein Zusammenhang mit der App-Nachfrage und der Zahlungsbereitschaft beobachtet werden. Auch in der nur teilweise vergleichbaren Studie von TNS Emnid und vzbv (2015) wurde kein entsprechender Zusammenhang festgestellt.[30]

4.5.5 Zusammenfassung

Zusammenfassend lässt sich aus den Ergebnissen der Befragung ableiten, dass Verbraucherinnen und Verbraucher ein Interesse an Datenschutz-Tools haben. Die untersuchten, potentiellen Funktionen der DATENSCHUTZscanner-Anwendung erhielten durchweg hohe Relevanzbewertungen, wobei eine bestimmte Varianz zwischen den einzelnen Funktionen beobachtet werden konnte.

[29] Ibid., Seite 10-11
[30] Ibid., Seite 10-11

4.6 Konfigurationsoptionen je Betriebssystem für Betroffene

Die Konfigurationsoptionen mit datenschutzrechtlicher Implikation sind für Betroffene nicht einheitlich: je nach Endgerät und insbesondere je nach Betriebssystem können Betroffene die Datenverarbeitungen unterschiedlich an die eigenen Bedürfnisse anpassen.

Selbst innerhalb eines Betriebssystems können sich die Optionen je Version oder Hersteller des mobilen Endgerätes unterscheiden. Hieraus ergeben sich aus Betroffenensicht mögliche Abweichungen zwischen Erwartung und Realität.

Eine App, die Betroffene bei der Wahrnehmung der eigenen Rechte unterstützt, sollte auf bestehende Optionen der mobilen Endgeräte zurückgreifen. Hierbei sollte eine solche App auch ein mögliches Delta zwischen Erwartung und Realität, bezogen auf die Konfigurationsoptionen der Betroffenen, adressieren. Eine technische und praktische Analyse der bestehenden Optionen je Betriebssystem wurde durchgeführt, um überhaupt im Rahmen des Forschungsvorhabens einschätzen zu können, welche möglichen Interessen seitens der Betroffenen bestehen könnten, und wie etwaige Machbarkeiten für Funktionalitäten einzuschätzen sind, die derartige Interessen abbilden. Diese Überlegungen sind auch in die Vorbereitung und Umsetzung der im Verlauf des Projektes durchgeführten Befragungen eingeflossen; so konnte vermieden werden, dass Interessen abgefragt werden, deren Umsetzung bereits frühzeitig als unrealistisch eingestuft wurden.

4.6.1 Generelles

Die bestehenden und möglichen Erwartungen der Betroffenen können auf drei Dimensionen reduziert werden:

a) Verarbeitung bestimmter Daten für das gesamte Endgerät konfigurieren
b) App-spezifisch Zugriffe auf Daten und/oder Funktionen konfigurieren
c) Datumsspezifisch die Verarbeitung konfigurieren

Die Verarbeitung bestimmter Informationen für das gesamte Endgerät zu konfigurieren (lit.a) umfasst all jene Fälle, in denen Betroffene bei der Nutzung eines Endgerätes nie die Verarbeitung bestimmter Daten wünschen. Hierbei kann es sich um unterschiedliche Aspekte handeln wie zum Beispiel den Standort, Kommunikations-IDs (Bluetooth, W-Lan), aber auch die Verarbeitung von Daten zu bestimmten Zwecken, zum Beispiel Werbezwecke.

App-spezifisch Zugriffe auf Daten und/oder Funktionen zu konfigurieren (lit. b) umfasst die Fälle, in denen Betroffene für einzelne Apps den Zugriff auf bestimmte Daten und Funktionen (z.B. Kontakte oder Standort) einschränken können.

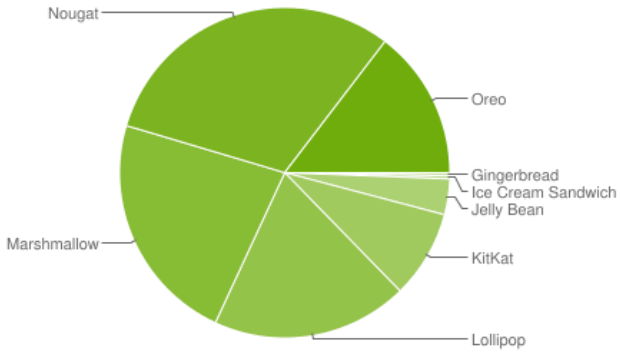
Datumsspezifisch die Verarbeitung zu konfigurieren (lit. c) umfasst jene Fälle, in denen der Zugriff nicht auf eine grundsätzliche Gruppe von Daten (Kontaktdaten) – egal ob wie in lit. a) für das gesamte Gerät oder wie in lit. b) nur für einzelne Apps – definiert werden kann. Vielmehr umfasst dies die Fälle, in denen bestimmte Einzeldaten aus dieser Gruppe definiert werden können: z.B. die Telefonnummer eines bestimmten Kontaktes.

4.6.2 Android

Android ist zum Teil ein open-source Betriebssystem[31]. Dies führt dazu, dass es eine vermeintliche Homogenität der Betriebssysteme – jedenfalls innerhalb der Betriebssystemversion – gibt. Faktisch bestehen zum Teil aber auch deutliche Unterschiede in den von den Geräteherstellern modifizierten Android-Versionen.

Im Rahmen des Forschungsprojektes galt es Lösungsansätze zu entwickeln, die diese Vielfalt adressieren. Das Forschungsprojekt konzentrierte sich hierbei auf die Android-Grundversion ohne Berücksichtigung etwaiger Modifikationen durch Endgerätehersteller. Zudem wurde hinsichtlich der Konfigurationsoptionen davon ausgegangen, dass Betroffene mindestens die Android-Version 6.0 (Marshmallow) verwenden (Tabelle 3: Verteilung der Android Versionen: Stand 08/2018

).



Version	Codename	API	Marktanteil
2.3.3 - 2.3.7	Gingerbread	10	0.3%
4.0.3 - 4.0.4	Ice Cream Sandwich	15	0.3%
4.1.x	Jelly Bean	16	1.1%
4.2.x		17	1.6%
4.3		18	0.5%
4.4	KitKat	19	7.8%
5.0	Lollipop	21	3.6%
5.1		22	14.7%
6.0	Marshmallow	23	21.6%
7.0	Nougat	24	19.0%
7.1		25	10.3%
8.0	Oreo	26	13.4%

Tabelle 3: Verteilung der Android Versionen: Stand 08/2018[32]

Die Wahl fiel auf Grund verschiedener Umstände auf diese Version. Einerseits war abzusehen, dass diese Version zum Ende des Forschungsprojektes eine sehr hohe Marktdurchdringung haben wird; andererseits sind ab Version 6.0 Betroffene unter Android in der Lage, Funktionszugriffe je App zu konfigurieren (vgl. Dimension 4.6.1 lit. b). Diese Funktionszugriffe können bezüglich der Funktionszugriffsgruppen Körpersensoren, Kalender, Kamera, Kontakte, Standort, Mikrofon, Telefon, SMS und Speicher

[31] https://source.android.com/
[32] https://developer.android.com/about/dashboards/

freigegeben oder entzogen werden. Da das Entziehen oder Freigeben derartiger Funktionszugriffe als ein elementarer Bestandteil der Selbstdatenschutzmechanismen des Forschungsprojektes eingestuft wurde, wurde dies als zwingende Voraussetzung für eine erfolgreiche Umsetzung der Projektergebnisse bei Betroffenen betrachtet und für die weitere Entwicklung diese Version als Minimum vorausgesetzt.

Hierbei gehen mit einer Funktionszugriffsgruppe oft mehrere Zugriffsrechte einher. Eine vollständige Auflistung der möglichen Zugriffsrechte ist den Entwicklerleitlinien zu entnehmen.[33] Eine Zuordnung der Zugriffsrechte zu den jeweiligen Funktionszugriffsgruppen ist ebenfalls den Entwicklerleitlinien zu entnehmen.[34] Nicht alle Zugriffsrechte sind einer Funktionszugriffsgruppe zugeordnet und können somit von Betroffenen beeinflusst werden.[35] Betroffene haben aber die Möglichkeit, die vollständige Liste der technisch eingeforderten Zugriffsrechte einzusehen.[36]

Permission Group	Permissions
CALENDAR	• READ_CALENDAR • WRITE_CALENDAR
CALL_LOG	• READ_CALL_LOG • WRITE_CALL_LOG • PROCESS_OUTGOING_CALLS
CAMERA	• CAMERA
CONTACTS	• READ_CONTACTS • WRITE_CONTACTS • GET_ACCOUNTS
LOCATION	• ACCESS_FINE_LOCATION • ACCESS_COARSE_LOCATION
MICROPHONE	• RECORD_AUDIO
PHONE	• READ_PHONE_STATE • READ_PHONE_NUMBERS • CALL_PHONE • ANSWER_PHONE_CALLS • ADD_VOICEMAIL • USE_SIP
SENSORS	• BODY_SENSORS
SMS	• SEND_SMS • RECEIVE_SMS • READ_SMS • RECEIVE_WAP_PUSH • RECEIVE_MMS
STORAGE	• READ_EXTERNAL_STORAGE • WRITE_EXTERNAL_STORAGE

Abbildung 11: Beispielhafte Darstellung der zugehörigen Rechte, die mit einer Funktionszugriffsgruppe einhergehen.[37]

[33] https://developer.android.com/reference/android/Manifest.permission
[34] https://developer.android.com/guide/topics/permissions/overview, Table 1.
[35] Vergleiche beispielsweise https://developer.android.com/guide/topics/permissions/overview#normal-dangerous; die dortige Unterscheidung der Berechtigungsklassen verdeutlicht, dass es neben den in Den Funktionszugriffsgruppen enthaltenen Berechtigungen auch weitere Berechtigungen gibt.
[36] So zum Beispiel durch Aufrufen von Einstellungen > Apps > Auswahl einer konkreten App > Berechtigungen > Alle Berechtigungen.
[37] https://developer.android.com/guide/topics/permissions/overview

Betroffene können Funktionszugriffsgruppen ausschließlich in ihrer Gesamtheit der jeweils darin enthaltenen Zugriffsrechte gewähren oder ablehnen. Innerhalb einer Funktionszugriffsgruppe besteht keine Möglichkeit, die gewährten Zugriffsrechte zu spezifizieren.

Beispiel: Betroffene, die einer App die Funktionszugriffsgruppe „Telefon“ erlauben, können – technisch – nicht weiter beeinflussen, ob diese App ausschließlich in der Lage ist zu erfahren, ob Betroffene aktuell „telefonieren“ – zum Beispiel um während eines Telefonats keine Benachrichtigungen zu senden – (Zugriffsrecht „READ_PHONE_STATE“) oder ob diese App auch das in der Funktionszugriffsgruppe enthaltene Zugriffsrecht „ANSWER_PHONE_CALLS“[38] – um selbständig Anrufe anzunehmen oder abzulehnen – nutzt, oder ob diese App eventuell über das ebenfalls enthaltene Zugriffsrecht „CALL_PHONE“[39] einen Anruft tätigt.[40]

Betroffenen ist es ebenfalls möglich, bestimmte Funktionen für das gesamte Gerät einzuschränken, so zum Beispiel den Standort oder die Konnektivität (Bluetooth, W-Lan). Nicht möglich ist es, für das Endgerät die Verarbeitung zu bestimmten Zwecken, wie zum Beispiel zu Werbezwecken, zu unterbinden. Allerdings bietet Android eine sogenannte Android-Werbe-ID. Diese Werbe-ID ist dynamisch und kann von Betroffenen jederzeit neugeneriert werden. Bei regelmäßiger und häufiger Iteration der Werbe-ID kann ein Rückschluss auf einzelne Endgeräte und Betroffene somit stark eingeschränkt werden.[41]

4.6.3 iOS

Auch iOS unterscheidet sich hinsichtlich der Funktionen zwischen den jeweiligen Versionen. Die bei Betroffenen in Verwendung befindlichen iOS-Versionen unterscheiden sich indessen deutlich weniger als die bei Android. Zudem bestehen keine Unterschiede innerhalb einer iOS-Version, da es keine unterschiedlichen Endgerätehersteller gibt. Darüber hinaus ist die Adaptionsrate bei verfügbaren iOS Updates wesentlich höher als bei anderen mobilen Betriebssystemen.

Insofern kann davon ausgegangen werden, dass Betroffene über eine aktuelle iOS Version verfügen, die eine Konfiguration der Zugriffsrechte ermöglicht.

iOS ermöglicht die Konfiguration der Zugriffsrechte je App. Hierbei wird ebenfalls auf Funktionszugriffsgruppen zurückgegriffen. Dies umfasst die Funktionszugriffsgruppen Contacts, Microphone, Calendars, Camera, Reminders, HomeKit, Photos, Health, Motion activity and fitness, Speech recognition, Location Services, Bluetooth sharing, Apple Music, media library, music and video activity, Social media accounts.[42] Auch hier ist lediglich der Zugriff auf eine Funktionszugriffsgruppe durch den Betroffenen beeinflussbar, ohne dies detaillierter konfigurieren zu können. Betroffene haben ebenfalls keinen Einfluss auf die Verwendungszwecke. iOS bietet entsprechend der Werbe-ID unter Android eine AD-Id.

[38] ANSWER_PHONE_CALLS: added in API level 26; Allows the app to answer an incoming phone call; https://developer.android.com/reference/android/Manifest.permission#ANSWER_PHONE_CALLS
[39] CALL_PHONE: added in API level 1; Allows an application to initiate a phone call without going through the Dialer user interface for the user to confirm the call; https://developer.android.com/reference/android/Manifest.permission#ANSWER_PHONE_CALLS
[40] https://developer.android.com/guide/topics/permissions/overview, Table 1.
[41] Vgl. hierzu auch https://developer.android.com/training/articles/user-data-ids; dort unter „Working with advertising IDs“.
[42] https://www.apple.com/business/site/docs/iOS_Security_Guide.pdf, Seite 73.

4.6.3.1 WindowsPhone

Das Microsoft Smartphone-Betriebssystem “WindowsPhone” war als designierter Nachfolger von Windows CE entwickelt worden. Es bestand eine hohe Wahrscheinlichkeit, dass zumindest im Unternehmensumfeld dieses Betriebssystem zur Anwendung kommt, da das lange Zeit vorherrschende “Blackberry OS” abgekündigt wurde. Zudem subventionierte Microsoft die Hardware in diesem Umfeld massiv. Jegliche Sicherheits- und Datenschutzfeatures waren unfertig und als “work in progress” zu verstehen. Das Rechte-Management war noch nicht standardisiert und die Entwicklungs-APIs noch in ständigem Wandel.

Im Rahmen des Forschungsprojektes wurde daher nur stichprobenartig auf die Umsetzbarkeit von Teilergebnissen im WindowsPhone-Kontext eingegangen. Es wurde evaluiert, inwiefern eine automatische Analyse der Installationsdateien möglich ist und ob die Erfahrungen mit der Verhaltensanalyse der anderen mobilen Betriebssysteme hier Anwendung finden können. Sowohl die Installationsdateienakquise als auch eine einfache Verhaltensanalyse wurden realisiert – bis Microsoft während der Laufzeit des Forschungsprojektes selbst WindowsPhone abgekündigt und die Weiterentwicklung eingestellt hat. Zu diesem Zeitpunkt wurde die weitere Projektarbeit auf die übrigen, etablierten Betriebssysteme konzentriert.

4.6.4 Zusammenfassung

Die Transparenz und Möglichkeit der Einflussnahme auf Datenverarbeitungen unterscheiden sich zum Teil erheblich für Nutzerinnen und Nutzer. Neben dem Betriebssystem spielen hierbei auch Modifikationen der Gerätehersteller eine Rolle. Im Forschungsprojekt wurde sich einerseits wegen der Verbreitung bei Verbraucherinnen und Verbrauchern, aber auch wegen der technischen Möglichkeiten auf eine Umsetzung unter Android konzentriert. Um eine Kernfunktion für dieses Forschungsprojekt abbilden zu können, wurde zudem mindestens die Version 6.0 vorausgesetzt. Einfluss können Betroffene allenfalls auf Ebene der Funktionszugriffsgruppen nehmen; datumsspezifische Konfigurationen sind über das Betriebssystem nicht vorgesehen; derartige Einschränkungsmöglichkeiten können aber im Einzelfall in einer App für diese App individuell angeboten werden.

4.7 Smartphonenutzung und Verbreitungsgrade der Betriebssysteme

Die möglichen Nutzungsszenarien und Funktionalitäten standen auch in Wechselwirkung zu den Befragungen. So bestätigten die Befragungen den aus technischer Sicht sinnvollen Fokus (Android ab Version 6.0). Sowohl in der Befragung 2016 als auch in der Befragung 2018 wurden die Teilnehmerinnen und Teilnehmer gebeten, Fragen zu ihrer Smartphone-Nutzung und ihrem Betriebssystem zu beantworten. Zuerst werden die Ergebnisse der ersten Befragung vorgestellt, hiernach die Ergebnisse der zweiten Befragung.

4.7.1 Smartphonenutzung und Verbreitungsgrade der Betriebssysteme in der Befragung 2016

4.7.1.1 Anzahl installierter Apps

90% der Teilnehmerinnen und Teilnehmer geben an auf ihrem aktuellen Smartphone bereits eine oder mehrere Apps selbst installiert zu haben. 8% der Befragten verneinen dies und 3% sind sich nicht sicher.

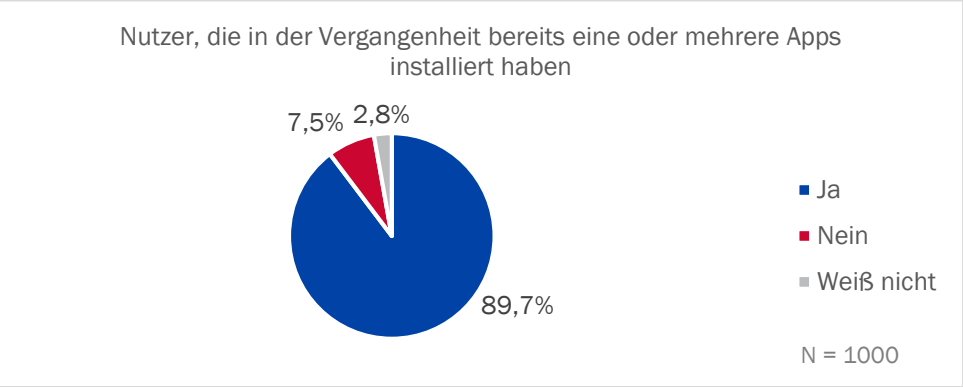


Abbildung 12: Anteil der Befragten, die App(s) installiert haben (Erste Befragung)

Die Anzahl der installierten Apps stellt sich wie folgt dar: 25% geben an, zwischen einer und fünf Apps geladen zu haben. 24% installierten sechs bis zehn Apps, 22% installierten elf bis 20 Apps, 10% 21 bis 30 Apps und 9% über 30 Apps.

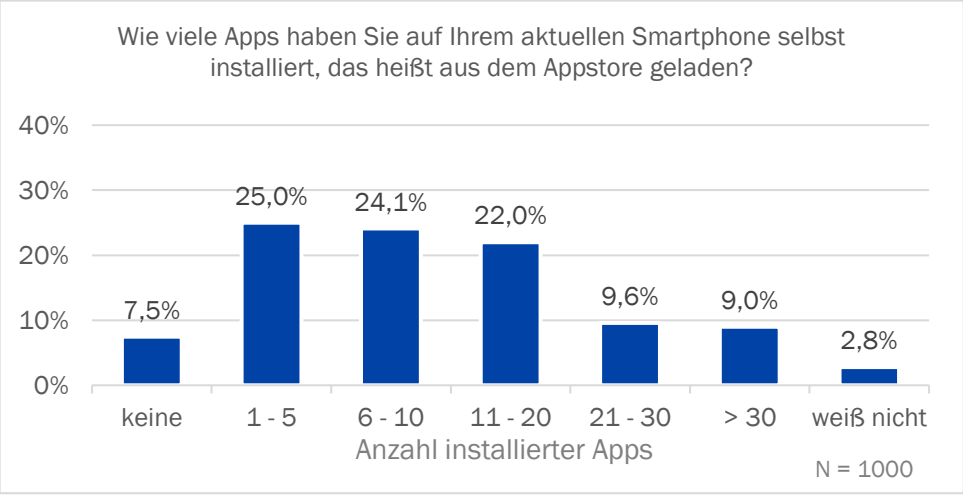


Abbildung 13: Anzahl selbst installierter Apps (Erste Befragung)

4.7.1.1.1 Unterschiede zwischen Geschlechtern und Altersgruppen

Männer und Frauen geben zu gleicher Häufigkeit an, in der Vergangenheit eine App auf ihrem Smartphone installiert zu haben. Zwischen der Anzahl der derzeitigen App-Installationen bei Frauen und Männern ist jedoch ein signifikanter Unterschied zu beobachten. So geben weibliche Teilnehmer an, dass sie mehr Apps auf ihrem Smartphone installiert haben als männliche Teilnehmer.

Weiterhin kann ein Alterseffekt bei der Anzahl der App-Installationen festgestellt werden. So steigt der Anteil an Teilnehmerinnen und Teilnehmern, die keine App auf ihrem Smartphone installiert haben mit dem Alter. Die Mehrheit der über 40-Jährigen hat zwischen einer und fünf Apps installiert, bei den 20 bis 39-Jährigen hat die Mehrheit zwischen sechs und zehn Apps installiert und bei den unter 20-Jährigen werden elf bis 20 installierte Apps am häufigsten genannt.

4.7.1.2 Appkäufe

Von den Teilnehmerinnen und Teilnehmern, die bereits eine oder mehrere Apps installiert haben (N=897), geben 35% an, dass sie für diese auch schon Geld bezahlt haben.

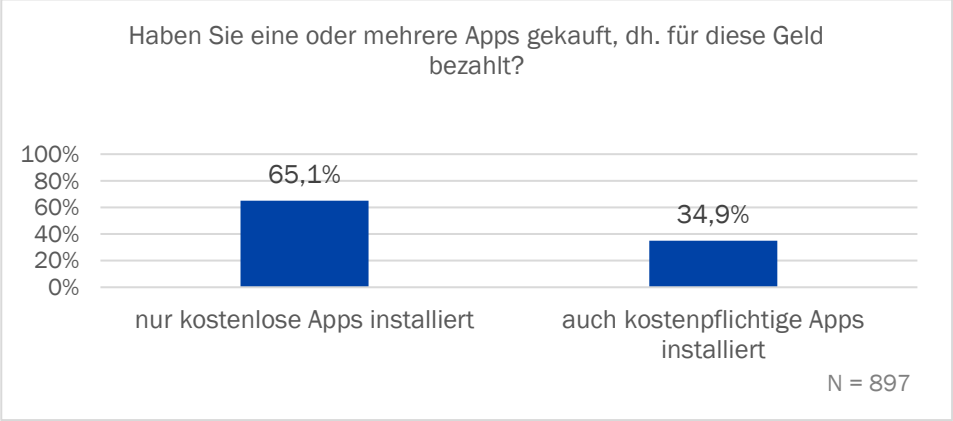


Abbildung 14: Anteil der Befragten, die kostenpflichtige Apps installiert haben (Erste Befragung)

4.7.1.2.1 Unterschiede zwischen Bevölkerungsgruppen

Bei App-Käufen ist ein starker signifikanter Effekt zwischen den Geschlechtern zu beobachten. 42% der Frauen geben an bereits für eine App Geld ausgegeben zu haben. Bei Männern liegt der Anteil bei 28%.

Auch zwischen den Altersklassen ist ein signifikanter Unterschied festzustellen. So fällt der Anteil von Smartphone-Nutzern, die Geld für eine oder mehrere Apps ausgegeben haben, mit dem Alter.

Für die verschiedenen Einkommenskategorien beobachtet man einen weiteren Effekt. So steigt mit dem monatlichen Nettoeinkommen auch der Anteil der Personen, die bereits für eine App Geld ausgegeben haben.

4.7.1.3 Betriebssystem

72% der Befragungsteilnehmer geben an, ein Android-Betriebssystem auf ihrem Smartphone installiert zu haben. 16% nutzen iOS, 8% WindowsPhone und 1% geben an ein anderes Betriebssystem zu nutzen. Weitere 3% sind sich nicht sicher, welches Betriebssystem auf ihrem Smartphone installiert ist.

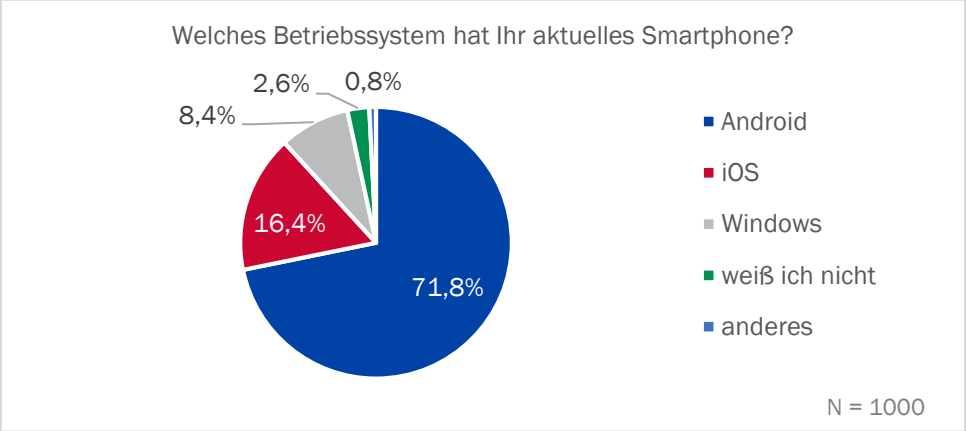


Abbildung 15: Betriebssystem des Smartphones (Erste Befragung)

4.7.2 Smartphonenutzung und Verbreitungsgrade der Betriebssysteme in der Befragung 2018

Bei der Anzahl der installierten Apps auf dem Smartphone ergeben sich in der zweiten Befragung ähnliche Werte wie in der ersten Befragung. Die Mehrheit der Befragten gibt an, zwischen einer und 20 Apps installiert zu haben. Hiervon haben 24% zwischen einer und fünf Apps installiert, 28% zwischen sechs und zehn Apps, 21% zwischen elf und 20 Apps. Weitere 11% geben an, dass sie zwischen 21 und 30 Apps installiert haben und weitere 8% besitzen mehr als 30 Apps. Der Anteil der Teilnehmerinnen und Teilnehmer, die angeben keine Apps installiert zu haben, ist im Vergleich zur ersten Befragung geringer und beträgt 5%. Weitere 4% der Befragten geben an nicht zu wissen, ob sie bisher eine App installiert haben.

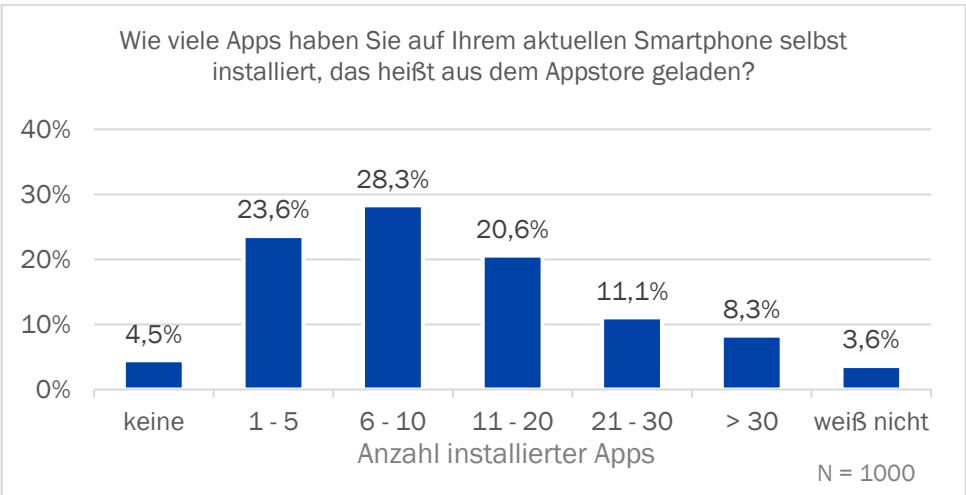


Abbildung 16: Anzahl selbst installierter Apps (Zweite Befragung)

4.7.2.1.1 Unterschiede zwischen Geschlechtern und Altersgruppen

Die Anzahl der installierten Apps unterscheidet sich zwischen Männern und Frauen, jedoch werden andere Ergebnisse beobachtet als in der ersten Befragung. So haben Männer signifikant mehr Apps auf ihrem Smartphone installiert als Frauen.

Die Ergebnisse zum Zusammenhang von App-Anzahl und Alter decken sich in der zweiten Befragung mit den Ergebnissen der ersten Befragung. So ist der Anteil an Teilnehmerinnen und Teilnehmern, die keine App auf ihrem Smartphone installiert haben, bei den älteren Gruppen höher. Die jüngeren Teilnehmerinnen und Teilnehmer (unter 30 Jahre) haben mehrheitlich elf oder mehr Apps installiert. Bei den über 30-Jährigen liegt der Durschnitt geringer und so geben in der Altersgruppe über 60 Jahre fast 2/3 der Befragten an, dass sie weniger als zehn Apps installiert haben.

4.7.2.2 App-Käufe

Von den Teilnehmerinnen und Teilnehmern, die angegeben haben, dass sie bereits eine App auf ihrem Smartphone installiert haben (N=955), geben 27% an, dass sie auch kostenpflichtige Apps installiert haben. Die restlichen 73% geben an, nur kostenlose Apps installiert zu haben. Im Vergleich zur ersten Befragung ist somit der Anteil der Teilnehmerinnen und Teilnehmer mit kostenpflichtigen Apps geringer als in der zweiten Befragung.

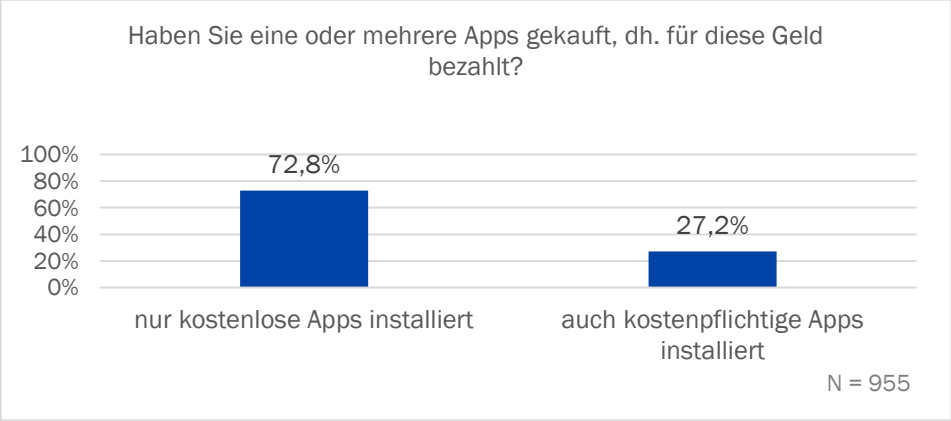


Abbildung 17: Anteil der Befragten, die kostenpflichtige Apps installiert haben (Zweite Befragung)

4.7.2.2.1 Unterschiede zwischen Bevölkerungsgruppen

Bei App-Käufen ist ein starker signifikanter Effekt zwischen den Geschlechtern zu beobachten. So geben 20% der Frauen an bereits für eine App Geld ausgegeben zu haben. Bei Männern liegt der Anteil bei 35%. Die Ergebnisse unterscheiden sich somit von denen der ersten Befragung.

Bezüglich des Alterseffekts können die Ergebnisse der ersten Befragung jedoch bestätigt werden. So wird in der zweiten Befragung ebenfalls ein signifikanter, jedoch schwacher Zusammenhang zwischen Alter und der Bereitschaft zu App-Käufen beobachtet.

Für die verschiedenen Einkommenskategorien beobachtet man lediglich einen schwachen Zusammenhang, der nicht auf einen Zusammenhang zwischen Einkommen und der Bereitschaft zu App-Käufen

schließen lässt. Dieses Ergebnis unterscheidet sich somit vom Ergebnis der ersten Befragung im Jahr 2016.

4.7.2.3 Betriebssystem

Die Ergebnisse der Verteilung der unterschiedlichen Smartphone-Betriebssysteme ist ähnlich zu denen in der ersten Befragung. Die große Mehrheit der Befragten, nämlich 74% geben an, ein Android-Smartphone zu nutzen. 20% nutzen iOS. Der Anteil der WindowsPhone-Nutzerinnen und Nutzer beträgt 4% und weniger als 1% der Befragten geben an, dass sie ein anderes Betriebssystem nutzen. Weitere 2% sind sich nicht sicher, welches Betriebssystem ihr Smartphone hat.

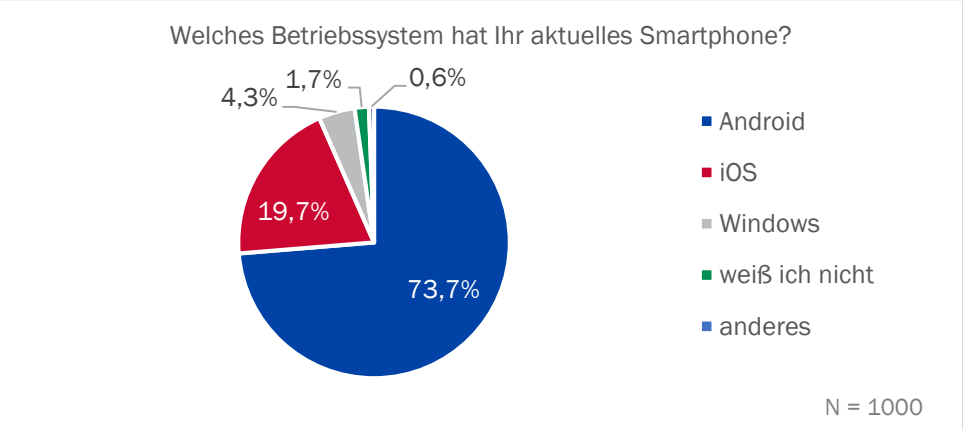


Abbildung 18: Betriebssystem des Smartphones (Zweite Befragung)

4.7.3 Zusammenfassung

Die Umfrageergebnisse bestätigen die bereits zu Beginn des Forschungsvorhabens als auch in der grauen Literatur genannten Marktanteile der unterschiedlichen Betriebssysteme[43]. Die sich aus den Umfragen für den deutschen Markt ergebenden Verbreitungsgrade galt es auch im Rahmen des Forschungsprojektes zu würdigen.

Zu Beginn des Forschungsprojekts stand auch die Entwicklung von Analysemethoden für Windows-Phone Apps im Fokus. Auf Basis der Verbreitungsgrade und der Ankündigung von Microsoft, Windows-Phone nicht weiter zu entwickeln, wurde die Erforschung und Entwicklung entsprechender Methoden eingestellt.

Hinsichtlich der beiden führenden Betriebssysteme Android und iOS galt es im Forschungsprojekt die Forschungsmittel zielführend einzusetzen. Vor dem Hintergrund der Verbreitungsgrade im Businesss-2-Consumer Markt wurde die Entwicklung entsprechender Methoden für Android priorisiert. Hierbei ist zu beachten, dass die technische Umsetzung und Forschung aufgrund des in weiten Teilen „offenen“ Betriebssystems ressourcen-effektiver umgesetzt werden kann. Die dabei gewonnen Erkenntnisse können aber in weiten Teilen durchaus auf iOS übertragen werden. Insbesondere im Bereich der Informations-

[43] So „Kantar: Apple gewinnt Marktanteile, Android verliert“, vom 8. Januar 2015, abzurufen unter https://www.heise.de/mac-and-i/meldung/Kantar-Apple-gewinnt-Marktanteile-Android-verliert-2513813.html.

aufbereitenden Algorithmen, siehe 4.11, ist eine nahezu vollständige Anschlussverwertung der für Android gewonnen Erkenntnisse möglich. Darüber hinaus sind die Erkenntnisse zu den Nutzeranforderungen und zur Bedienbarkeit des App-Clients, siehe 5.2, auf andere Betriebssysteme übertragbar. Lediglich im Rahmen der Generierung von hinreichend zuverlässigen Analyse-Rohdaten, siehe 4.9, und der Umsetzung bestimmter Funktionalitäten in einem zu entwickelnden iOS App-Client, siehe 5.1, wurde mit relevanten Abweichungen gerechnet.

Auf dieser Basis wurde zudem der sog. Demonstrator für Android entwickelt; Umsetzungsoptionen für iOS jedoch konzeptionell erforscht.[44]

4.8 Zielsetzung der Anwendungen und Informationstexte

4.8.1 Mission und Vision

Die Verbraucherbefragung zeigte, dass bei Smartphone-Nutzerinnen und -Nutzern ein Bedarf an Datenschutz-Lösungen für das Smartphone existiert. Zum einen schätzen Verbraucherinnen und Verbraucher ihre Kompetenz lediglich mittelmäßig ein, was sich auch in ihrer tatsächlichen Datenschutzkompetenz widerspiegelt, zum anderen haben sie eine gewisse Motivation, etwas für ihren Datenschutz zu tun. Die beschriebene Nachfrage nach einer unterstützenden Anwendung, wie dem DATENSCHUTZscanner, bekräftigt den Bedarf an einer Lösung.

Um die Zielsetzung des Vorhabens zu schärfen, wurde deshalb das Mission- und Vision-Statement der DATENSCHUTZscanner-Anwendungen erarbeitet. Dieses Statement stellt die Leitplanken für die relevanten Inhalte und Funktionen der Anwendungen dar.

4.8.1.1 Methodik

Die Entwicklung des Statements basierte methodisch auf einem Konsortialworkshop, der im März 2017 stattfand. Zur Vorbereitung wurden die Ergebnisse der Literaturanalyse aus dem Bereich der Verhaltenswissenschaften kondensiert und fünf relevante Analysedimensionen definiert. Diese sind:

1. Pain: Was ist das Problem der Nutzerinnen und Nutzer?
2. Gain: Was können potentielle Nutzerinnen und Nutzer durch die Anwendung erreichen?
3. Value: Was ist der Mehrwert der Anwendung in der Zukunft?
4. Team: Wer, das heißt welches Team mit seinen konkreten Kompetenzen, erarbeitet die Anwendung?

Die letzte Analysedimension leitet sich aus den obigen ab und fasst den zentralen Anspruch zusammen:

5. Solution: Wie sieht die Lösung aus? Wie kann die Anwendungsmission beschrieben werden?

Während des Workshops wurden die fünf Analysedimensionen inhaltlich ausgefüllt und konkretisiert.

4.8.1.2 Ergebnisse

Die Ergebnisse in den fünf Analysedimensionen können wie folgt zusammengefasst werden.

[44] Siehe 5.

4.8.1.2.1 Analysedimension 1: Pain

Basierend auf den Ergebnissen der ersten Befragung und der verhaltenswissenschaftlichen Literatur zum Thema Datenschutz kann festgestellt werden, dass Verbraucherinnen und Verbraucher ein besonderes Unwohlsein verspüren. Aufgrund der technischen und juristischen Komplexität des Themas existiert darüber hinaus eine gewisse Überforderung.[45] Da der Handlungsrahmen für Verbraucherinnen und Verbraucher derzeit bei der Ausübung ihrer informationellen Selbstbestimmung beschränkt ist, kann man zusätzlich eine gewisse Alternativlosigkeit feststellen.[46]

4.8.1.2.2 Analysedimension 2: Gain

Um die genannten Schwierigkeiten der Verbraucherinnen und Verbraucher zu adressieren, soll mit den DATENSCHUTZscanner-Anwendungen eine Lösung geschaffen werden, die das Unwohlsein, die Überforderung und Alternativlosigkeit auflöst. Übergeordnet bedeutet das, dass Verbraucherinnen und Verbraucher mit den DATENSCHUTZscanner-Anwendungen selbstbestimmt handeln können.

4.8.1.2.3 Analysedimension 3: Value

Der zukünftige Mehrwert erfüllt das Selbstbestimmungs-Ziel und soll neben möglichst vollständigen, präzisen und objektiven Informationen dafür sorgen, dass Nutzerinnern und Nutzer auf Basis dieser Transparenz bewusster entscheiden und sich somit auf dem App-Markt auch funktionsgleiche, aber datensparsamere Alternativen entwickeln können, das heißt neue Alternativen durch veränderten Bedarf des Marktes.

4.8.1.2.4 Analysedimension 4: Team

Um all diese Ziele zu erreichen, ist ein Team mit unterschiedlichen Kompetenzen notwendig. Dieses zeichnet sich durch Vielseitigkeit aus und hat aus technischer, juristischer und verhaltenswissenschaftlicher Perspektive die Fähigkeiten, Nutzerinnen und Nutzer sowie ihre Bedürfnisse zu verstehen und diesbezüglich Lösungen zu entwickeln.

4.8.1.2.5 Analysedimension 5: Solution

Aus diesen Ergebnissen leitet sich die Mission und Vision der DATENSCHUTZscanner-Anwendungen ab.

Der DATENSCHUTZscanner bietet Nutzerinnen und Nutzern mehr Transparenz und Kontrolle beim Thema Datenschutz in ihren Apps.

4.8.2 Informationstexte

Die Informationstexte sind Kernbestandteil der Forschungsarbeiten und sowohl in der App, als auch in den übrigen Labormustern enthalten. Sie sollen Nutzerinnen und Nutzer knapp und präzise über Datenverarbeitungen informieren und so das Transparenzziel der Anwendungen sicherstellen. Darüber hinaus beinhalten die Informationstexte Erläuterungen zu Handlungsoptionen, die Nutzerinnen und Nutzer umsetzen können, um eine bestimmte Datenverarbeitung zu beheben oder zu vermeiden.

[45] Der Aspekt der Überforderung wurde ebenfalls von den Nutzerinnen und Nutzern auf der Internationalen Funkausstellung Berlin im September 2017 bestätigt und wurde von den Teilnehmerinnen und Teilnehmern in den Fokusgruppen im Januar 2018 betont.
[46] Vgl. Kettner, S.E., Thorun, C. & Vetter, M. (2018). Wege zur besseren Informiertheit: Verhaltenswissenschaftliche Ergebnisse zur Wirksamkeit des One-Pager-Ansatzes und weiterer Lösungsansätze im Datenschutz, S. 77.

4.8.2.1 Aufbau der Informationstexte

Die Informationstexte folgen in ihrem Aufbau einem festgelegten Schema. Jeder Informationstext besteht aus einer **übergeordneten Überschrift**, die grob beschreibt, welche Datenverarbeitung vorliegt.

Zusätzlich beinhaltet der Informationstext eine **Erläuterung der Konsequenzen** der Datenverarbeitung. Diese können zum einen positiv sein, in Form von potentiellen Vorteilen aus der Datenverarbeitung. Auf der anderen Seite werden negative Konsequenzen dargestellt, die aufzeigen, welche potentiellen Nachteile aus der Datenverarbeitung resultieren.

Auf diese Art und Weise werden Nutzerinnen und Nutzer nicht nur einseitig und „schwarzmalerisch" über Datenverarbeitungen informiert, sondern sie erhalten einen vollständigen Überblick über die Folgen einer Datenverarbeitung und können sich ihr eigenes Bild machen.

Beispiel 1: Überschrift und Erläuterungen der Konsequenzen

Überschrift	Die Datenschutzerklärung verwendet ungenaue Formulierungen	
Erläuterungen der Konsequenzen	⊕ Die Datenschutzerklärung kann hierdurch kürzer und übersichtlicher werden.	⊖ Der Anbieter kann den Umfang der Datenverarbeitungen so verändern, ohne dass Sie die Änderungen im Text erkennen können. ⊖ Der Anbieter kommt seiner Aufklärungspflicht nur teilweise nach, sodass Sie sich nicht hinreichend über die Datenverarbeitungen informieren können.

Bei Bedarf wurden sowohl für die Überschrift als auch für die Erläuterungen der Konsequenzen **zusätzliche Erklärungen und Detailinformationen** erstellt. Diese befinden sich auf der „zweiten Unterebene“ des Informationstextes.

Beispiel 2: Beispiel 1 + zweite Unterebene mit Detailinformationen

Überschrift	Die Datenschutzerklärung verwendet ungenaue Formulierungen	
Detailinformation zur Überschrift	Die Datenschutzerklärung nutzt Formulierungen wie "z.B.", "in der Regel" oder "grundsätzlich". Sie können sich nicht sicher sein, welche und wie Ihre Daten tatsächlich durch die App verarbeitet werden.	
Erläuterungen der Konsequenzen	⊕ Die Datenschutzerklärung kann hierdurch kürzer und übersichtlicher werden.	⊖ Der Anbieter kann den Umfang der Datenverarbeitungen so verändern, ohne dass Sie die Änderungen im Text erkennen können. ⊖ Der Anbieter kommt seiner Aufklärungspflicht nur teilweise nach, sodass Sie sich nicht hinreichend über die Datenverarbeitungen informieren können.
Detailinformationen zur Erläuterung der Konsequenzen		⊖ Die Datenschutzerklärung ist zulässig, soweit die Änderungen geringfügig sind. Erhebliche Änderungen müssen Ihnen mitgeteilt werden.

Über die Erläuterungen der Datenverarbeitungen hinaus, zeigen die Informationstexte auch **Handlungsoptionen** auf. Diese Handlungsoptionen beschreiben, welche Maßnahmen Nutzerinnen und Nutzer ergreifen können, um eine Datenverarbeitung zu vermeiden. Darüber hinaus existieren für Teile der Informationstexte zusätzliche Informationen auf einer zweiten Unterebene, die bei Bedarf angezeigt werden können.

Beispiel 3: Beispiel 2 + Handlungsempfehlungen

Überschrift	Die Datenschutzerklärung verwendet ungenaue Formulierungen	
Detailinformation zur Überschrift	Die Datenschutzerklärung nutzt Formulierungen wie "z.B.", "in der Regel" oder "grundsätzlich". Sie können sich nicht sicher sein, welche und wie Ihre Daten tatsächlich durch die App verarbeitet werden.	
Erläuterungen der Konsequenzen	⊕ Die Datenschutzerklärung kann hierdurch kürzer und übersichtlicher werden.	⊖ Der Anbieter kann den Umfang der Datenverarbeitungen so verändern, ohne dass Sie die Änderungen im Text erkennen können. ⊖ Der Anbieter kommt seiner Aufklärungspflicht nur teilweise nach, sodass Sie sich nicht hinreichend über die Datenverarbeitungen informieren können.
Detailinformationen zur Erläuterung der Konsequenzen		⊖ Die ist zulässig, soweit die Änderungen geringfügig sind. Erhebliche Änderungen müssen Ihnen mitgeteilt werden.
Handlungsoptionen	Wir haben (X) relevante Textabschnitte gefunden. Unsere Empfehlung(en) Lesen Sie diese Textabschnitte und entscheiden Sie selbst, ob Sie die App unter diesen Umständen weiter verwenden möchten. Wenn Sie unsicher sind, ob die Datenverarbeitung Ihren Vorstellungen entspricht, verbieten Sie der App alle Zugriffe auf Daten, die aus Ihrer Sicht besonders schützenswert sind (z.B. Fotos, Adressbuch, Standort oder Kalender). Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!	

4.8.2.2 Rote Linien

Grundsätzlich sind die Informationstexte neutral gestaltet. Nutzerinnen und Nutzer erhalten keine Vorbewertung, sondern können mit Hilfe der detaillierten Informationen selbst entscheiden, wie sie eine Datenverarbeitung bewerten.

Da jedoch Teile der Informationstexte auf Datenverarbeitungen hinweisen, bei denen ein Gesetzesverstoß vorliegt, werden diese besonderen Informationstexte speziell hervorgehoben und als sogenannte „rote Linien“ definiert. Bei diesen roten Linien wird Nutzerinnen und Nutzern empfohlen, die entspre-

chende App nicht zu benutzen oder andere Maßnahmen zum Selbstdatenschutz anzuwenden. Nutzerinnern und Nutzer können jedoch auch bei „roten Linien“, den Fund als für sich selbst „unkritisch“ einstufen.

4.8.2.3 Cluster

Insgesamt wurden die Informationstexte sieben verschiedenen Clustern zugeordnet. Diese sind:

1. Datenschutzerklärungen, das heißt Aspekte, die die Datenschutzerklärungen der entsprechenden Anwendungen betreffen.
2. Datensicherheit, das heißt Aspekte die die Absicherung der Daten betreffen. Hierzu zählen Malware oder mangelhafte Verschlüsselungen.
3. Identifikation, das heißt Aspekte, die die Möglichkeit betreffen, ein Gerät durch den Einsatz digitaler Technologien bei der wiederholten Nutzung und Datenverarbeitung wiederzuerkennen.
4. Zugriffe, das heißt Aspekte, die das Recht auf Funktionen wie den Standort oder das Adressbuch zuzugreifen, betreffen.
5. Profilbildung, das heißt Aspekte, die es ermöglichen, aus den angegebenen und verarbeiteten Informationen ein Profil zu erstellen.
6. Werbung, das heißt Aspekte bei denen Werbung durch Drittanbieter ausgespielt oder aufgrund der Datenverarbeitung auf eine Person oder ein Gerät zugeschnitten wird.
7. Übertragung an Dritte, das heißt Aspekte, die die Weitergabe von verarbeiteten Daten an Dritte betreffen.

4.8.3 Erstellung, Überprüfung und Überarbeitung der Informationstexte

Im Rahmen der Erstellung der Informationstexte wurden im Projektzeitraum sechs Schritte absolviert.

1. Schritt: Erstellung erster Entwürfe für Informationstexte
2. Schritt: Überprüfung der Informationstexte durch Online-Befragung
3. Schritt: Überarbeitung der Informationstexte
4. Schritt: Überprüfung der Informationstexte in Fokusgruppen
5. Schritt: Finalisierung der Informationstexte
6. Schritt: Überprüfung der Informationstexte-Cluster durch Online-Befragung

Diese werden im Folgenden methodisch beschrieben und die zugehörigen Ergebnisse präsentiert.

4.8.3.1 Schritt 1: Erstellung erster Entwürfe für Informationstexte

Bis Juni 2016 wurden die ersten Entwürfe der Informationstexte erstellt. Die erste Auswahl der relevanten Aspekte basierte auf einer Auswertung der wissenschaftlichen und grauen Literatur. Diese Sammlung wurde ergänzt um diejenigen Aspekte, die aufgrund einer Analyse ausgewählter Datenschutzerklärungen von Apps für möglicherweise interessant erachtet wurden.

Bei der Auswahl der Datenschutzerklärungen zur ergänzenden Analyse wurde darauf geachtet, dass die Datenschutzerklärungen die Situation der Betroffenen gut abbilden. Bei der Auswahl wurden folgende Attribute berücksichtigt:

a) große Verbreitung,
b) kommerzielle und nicht kommerzielle Angebote,
c) unterschiedliche Kategorien (Spiel, Shopping, Entertainment, Nachrichten, usw.),
d) Angebote von öffentlichen und nicht-öffentlichen Stellen.

Datenerhebungen, die bei dieser Untersuchung wiederholt auftraten bzw. identifiziert wurden, wurden in einen Katalog aufgenommen und zur weiteren Bearbeitung vorbereitet.

4.8.3.2 Schritt 2: Überprüfung der Informationstexte durch Online-Befragung

Die Informationstexte wurden sodann einer ersten Überprüfung mit potentiellen Nutzerinnen und Nutzern unterzogen. Ziel der Überprüfung war es, herauszufinden, welche Relevanz die erstellten Informationstexte für Nutzerinnen und Nutzer haben. Die Methodik und die Ergebnisse der Überprüfung werden im Folgenden beschrieben.

4.8.3.2.1 Methodik: Überprüfung der Informationstexte durch Online-Befragung

Die erste Überprüfung der Informationstexte wurde mithilfe einer Online-Befragung durchgeführt. Die übergeordnete methodische Beschreibung findet sich in 4.2. Da es das Ziel der Befragung war, herauszufinden, welche Relevanz die erstellten Informationstexte für Nutzerinnen und Nutzer haben, wurden die 26 Informationstexte angezeigt und anhand der folgenden Fragestellung getestet:

Falls Ihr Smartphone von dieser Datenverarbeitung betroffen ist oder eine ähnlich ungenügende Datenschutzerklärung vorliegt, wie sehr interessiert Sie dann ein Hinweis darüber?

Interessiert mich überhaupt nicht o o o o o o o *interessiert mich sehr*

Somit liegen die Antworten auf einer 7er-Skala mit 1 als geringsten und 7 als höchsten Wert.

4.8.3.2.2 Ergebnisse: Erste Überprüfung der Informationstexte durch Online-Befragung

Insgesamt lässt sich feststellen, dass alle 26 Informationstexte für die Befragten ein gewisses Informationsinteresse hervorrufen. Zwischen den einzelnen Informationstexten existiert eine gewisse Varianz in der Interessenslage, jedoch lässt sich als übergeordnetes Ergebnis aus der Befragung ableiten, dass die identifizierten Datenverarbeitungen, die sich in den Informationstexten wiederfinden, relevant sind und weiterverfolgt werden sollten.

Die folgenden Abbildungen zeigen sowohl die Anteile der Nutzerinnen und Nutzer, die einen entsprechenden Informationstext mit mindestens „ein wenig interessant“ bewertet haben (Abbildung 19), als auch die Durchschnittsbewertungen der einzelnen Informationsblöcke (Abbildung 20).

Der Anteil der Nutzerinnen und Nutzer, die bei der Bewertung mindestens “interessiert mich ein wenig“ geantwortet haben, liegt für alle Informationstexte zwischen 62% und 78%. Malware (78%), eine unsicher implementierte Verschlüsselung (77%) und die Übermittlung unverschlüsselter Login-Daten (77%) erhalten die höchsten Interessenswerte. Werbenetzwerke (62%), Standortdaten (63%) und individualisierte Werbung durch Dritte und Partner (67%) werden als weniger interessant bewertet.

Ein ähnliches Bild zeichnet sich auch für die Durchschnittsbewertungen der Informationstexte ab. Diese liegen zwischen 4,94 und 5,75 mit einem globalen Mittelwert über alle Befragten und alle Informationsblöcke von 5,37 und einem Median von 5,69. Die Reihenfolge ist im Vergleich zur vorherigen Bewertung (Abbildung 19) minimal unterschiedlich. Malware (5,75), Datenverarbeitungen, die ausgeschlossen wurden (5,68) und die unverschlüsselte Übermittlung von Login-Daten (5,61) sind für die Teilnehmerinnen und Teilnehmer im Durchschnitt am interessantesten. Das heißt, dass ein Hinweis über eine App, die diese Praxis an den Tag legt, bei Nutzerinnen und Nutzern besonders gefragt ist. Werbenetzwerke (4,94), Standortnutzung (4,97) und das Erheben statischer Gerätekennungen (5,14) werden mit geringeren Durchschnittswerten versehen.

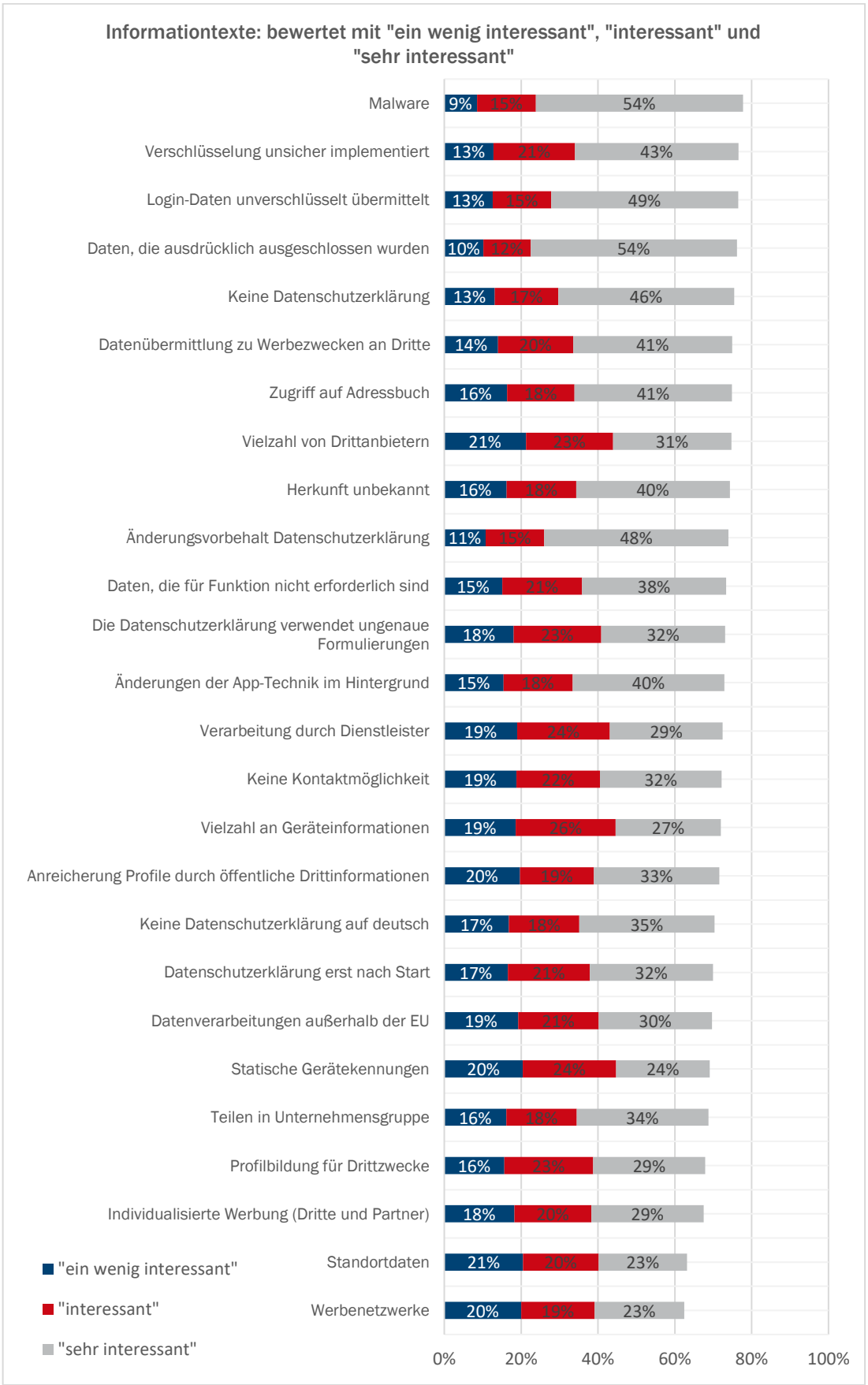


Abbildung 19: Informationstexte (Interessenskategorie)

Abbildung 20: Bewertung der einzelnen Informationstexte (Mittelwert)

4.8.3.2.2.1 Informationstexte und Unterschiede zwischen den Bevölkerungsgruppen

Zusätzlich zur übergeordneten Auswertung der Informationsexte wurden die Ergebnisse hinsichtlich potentieller Unterschiede bei Geschlecht, Alter, Bildung und Einkommen getestet. Die Ergebnisse können wie folgt zusammengefasst werden:

Geschlecht

Zwischen Männern und Frauen kann kein systematischer Unterschied in der Bewertung der Informationstexte festgestellt werden.

Alter

Die Bewertung der Informationstexte ist signifikant positiv, jedoch nur schwach mit dem Alter in Jahren korreliert (ρ=0,15, p<0,001). So beurteilen ältere Befragte die Informationstexte im Durchschnitt als interessanter als jüngere Befragte.

Bildung und Einkommen

Auch zwischen den Bildungsgruppen und Einkommensklassen kann kein Unterschied über die Bewertung der Informationstexte beobachtet werden.

4.8.3.3 Schritt 3: Überarbeitung der Informationstexte

Die Informationstexte wurden im nächsten Schritt überarbeitet und präzisiert. Einige Informationstexte wurden hinzugefügt, weitere wurden nicht weiterverfolgt, da sie technisch nicht analysierbar sind und somit keinen Mehrwert für Verbraucherinnen und Verbraucher in einer App oder Webseitenanwendung bieten.

Hierbei sind neben den Überlegungen aus der Verhaltensökonomie und der Sicherstellung der Verständlichkeit insbesondere auch Aspekte eingeflossen, die etwaige juristische Risiken eines Angebots, wie im Demonstrator realisiert, begründen könnten (siehe 5.5).

Bezüglich Tatsachenbehauptungen wurde regelmäßig die technische Entwicklung evaluiert und die Zuverlässigkeit und Reproduzierbarkeit der Ergebnisse überprüft. Als Tatsachenbehauptungen wurde im Projekt im Wesentlichen die Anzeige eines Prüfergebnisses zu einer bestimmten App eingestuft. Ergänzend wurden die Texte stets bezüglich möglicher Missverständnisse evaluiert und – soweit erforderlich – sprachlich geschärft.

Soweit die Forschung und Entwicklung im Rahmen der technischen und semantischen Analyse für die Generierung der Prüfergebnisse keine reproduzierbaren Ergebnisse lieferte, wurden die Informationstexte für die weitere Verarbeitung gesperrt, sodass diese Nutzerinnern und Nutzern nicht mehr angezeigt würden. Eine Löschung wurde nicht vorgenommen, um die bis dahin erreichten Ergebnisse nicht zu verlieren, soweit davon ausgegangen wurde, dass die weitere Forschung die notwendige Zuverlässigkeit erreichen könnte. Soweit die Prüfergebnisse zwar reproduzierbar waren, aber nicht mit einer hinreichenden Wahrscheinlichkeit oder aus anderen Gründen als der bloßen Reproduzierbarkeit Un-

schärfen aufwiesen, wurden die Informationstexte – soweit erforderlich – umformuliert. Im Wesentlichen wurden ursprünglich absolute Formulierung mit Relativierungen ergänzt. Derartige Relativierungen wurden einer Sperre der betroffenen Informationstexte immer dann bevorzugt, wenn die Forschungsergebnisse der sonstigen Disziplinen, insbesondere der Verhaltensökonomie, einen entsprechenden Bedarf der Aufklärung bei den Nutzerinnen und Nutzern, kombiniert mit einem entsprechenden Interesse bei den Nutzerinnen und Nutzern, feststellen konnte.

Ungeachtet der juristischen Aspekte fanden für jeden Informationstext regelmäßige Überprüfungen hinsichtlich des Mehrwerts für Nutzerinnern und Nutzer statt. Ziel des Forschungsprojektes ist es insbesondere, neben der bloßen Transparenz, Nutzerinnen und Nutzern sachdienliche Handlungsempfehlungen bereitzustellen. Soweit die technische Forschung für diese Empfehlungen Grenzen ermittelt hat, wurde der jeweilige Informationstext ebenfalls hinterfragt und – wenn aus Sicht des Konsortiums - ein Mehrwert für Betroffene mangels sachdienlicher Empfehlung nicht mehr erreicht werden konnte, wurde der entsprechende Informationstext ebenfalls gesperrt.

4.8.3.4 Schritt 4: Überprüfung der Informationstexte in Fokusgruppen

Im vierten Schritt wurden die aktualisierten Informationstexte einer weiteren Überprüfung an potentiellen Nutzerinnen und Nutzern unterzogen. Hierzu wurden im Januar 2018 drei Fokusgruppengespräche durchgeführt, deren Ziel es war, die Verständlichkeit zu untersuchen und weitere Änderungsnotwendigkeiten für die Finalisierung der Informationstexte aufzudecken.

4.8.3.4.1 Methodik: Überprüfung der Informationstexte in Fokusgruppen

Insgesamt wurden drei Gruppengespräche mit jeweils zehn Teilnehmerinnen und Teilnehmern terminiert und von den Mitarbeiterinnen und Mitarbeitern des Forschungsprojekts durchgeführt, transkribiert und ausgewertet. Die Rekrutierung erfolgte durch ein externes Marktforschungsinstitut. Tabelle 10 im Appendix (0) gibt einen Überblick über die Zusammensetzung der Gruppen.

4.8.3.4.1.1 Ablauf der Gruppengespräche

Die Gespräche dauerten jeweils zwei Stunden und folgten dem gleichen Ablaufplan.

Zuerst wurde eine kurze Einführung in das Forschungsprojekt und die Zielsetzung der Gespräche gegeben. Hiernach wurde eine kurze Vorstellungsrunde durchgeführt und die Informationstexte wurden vorgestellt. Um den Teilnehmerinnen und Teilnehmern die Inhalte und den Aufbau der Informationstexte plastisch vor Augen zu führen, wurden die Informationstexte „Malware“ und „Werbenetzwerke“ verwendet. Diese beiden wurden ausgewählt, da „Malware“ die höchste und „Werbenetzwerke“ die niedrigste Bewertung in der ersten Befragung (vgl. Abbildung 20) erhielten und somit den Relevanzrahmen der Informationstexte abgrenzten.

Der Hauptteil der Gespräche bestand aus der Diskussion der Informationstexte. Die zu behandelnden Informationstexte wurden in Abhängigkeit von den „Clustern“ zufällig einer der Gruppen zugeteilt. In der ersten und zweiten Gruppe wurden insgesamt neun Informationstexte diskutiert. In der dritten Gruppe wurden sieben Informationstexte diskutiert.

Die Teilnehmerinnen und Teilnehmer wurden gebeten, die Informationstexte einzeln kritisch zu reflektieren, Schwierigkeiten in der Verständlichkeit aufzudecken und Vorschläge zur Verbesserung dieser zu machen.

Die Untersuchung der Informationstexte beschränkte sich auf die erste Ebene der Informationstexte, inklusive positiver und negativer Konsequenzen der Datenverarbeitung und Handlungsempfehlungen.[47]

4.8.3.4.2 Ergebnisse: Überprüfung der Informationstexte in Fokusgruppen

Die konkreten Anmerkungen der Teilnehmerinnen und Teilnehmer zu den einzelnen Informationstexten sind in die finale Version der Informationstexte[48] eingeflossen und werden nicht einzeln aufgelistet. Übergeordnet ließ sich jedoch feststellen, dass die Verständlichkeit der Texte bereits sehr hoch war und insgesamt lediglich kleinere Textverbesserungen vorgenommen werden mussten. Dies betraf beispielsweise juristische und technische Fachbegriffe, bei denen sich die Teilnehmerinnen und Teilnehmer eine bessere Erläuterung der Datenverarbeitungen wünschten. Außerdem wünschten sich die Teilnehmerinnen und Teilnehmer Erklärungen für verwendete Fremdworte. Ein weiteres Ergebnis der Fokusgruppen war außerdem der Wunsch der Teilnehmerinnen und Teilnehmer, Warnungen in den Informationstexten expliziter auszusprechen und eine bessere Orientierungshilfe zu liefern. Dieser Aspekt wurde bei der Definition der "roten Linien" (vgl. 4.8.2.2) berücksichtigt und umgesetzt.

4.8.3.5 Schritt 5: Finalisierung der Informationstexte

Basierend auf den Ergebnissen der Fokusgruppen wurden die Informationstexte im Forschungsprojekt final geschärft. Eine Liste aller Informationstexte findet sich im Appendix, Punkt 9.1.

4.8.3.6 Schritt 6: Überprüfung der Informationstexte-Cluster durch Online-Befragung

Im letzten Schritt wurde die Relevanz der Cluster für Nutzerinnen und Nutzer in einer zweiten Online-Befragung überprüft. Ziel der Befragung war es, Forschungsbedarfe der einzelnen Cluster zu identifizieren und zu erörtern, bei welchen Aspekten / Clustern Nutzerinnen und Nutzer Schwierigkeiten haben, eine Datenverarbeitung zu erkennen.

4.8.3.6.1.1 Methodik: Überprüfung der Informationstexte-Cluster durch Online-Befragung

Übergeordnete Informationen zur Methodik der Befragung sind in 4.2 zusammenfasst. Zur Untersuchung der Forschungsfrage wurden zu den sieben Clustern der Informationstexte jeweils zwei Fragen gestellt:

a) Wie einfach oder schwierig ist es für Sie zu erkennen, dass „Aspekt / Cluster“ zutrifft.
b) Wie wichtig oder unwichtig ist es für Sie zu erkennen, dass „Aspekt / Cluster“ zutrifft.

Durch die erste Teilfrage sollte identifiziert werden, ob die Verbraucherinnen und Verbraucher in der Lage sind, einen bestimmten Datenschutzaspekt zu identifizieren und ob sie es einfach oder schwierig

[47] Diese Beschränkung wurde vorgenommen, da in der Anwendung lediglich die erste Ebene per Default angezeigt wird und somit festgestellt werden konnte, an welchen Stellen ohne zusätzliche Erklärung für die Teilnehmerinnen und Teilnehmer die Verständlichkeit nicht gewährleistet ist.
[48] vgl. Schritt 5 „Finalisierung der Informationstexte“

finden, diesen zu erkennen. Die zweite Teilfrage zielte darauf ab, den Bedarf der Verbraucherinnen und Verbraucher zu untersuchen, Informationen zu einem bestimmten Datenschutzaspekt zu erhalten.

In der Befragung wurden die folgenden sieben Cluster in acht Teilfragen abgedeckt:

1. Datenschutzerklärungen, das heißt Aspekte, die die Datenschutzerklärungen der entsprechenden Anwendungen betreffen.
2. Datensicherheit, das heißt Aspekte die die Absicherung der Daten betreffen. Hierzu zählen Malware oder mangelhafte Verschlüsselungen.
3. Identifikation, das heißt Aspekte, die die Möglichkeit betreffen, ein Gerät durch den Einsatz digitaler Technologien bei der wiederholten Nutzung und Datenverarbeitung wiederzuerkennen.
4. Zugriffe, das heißt Aspekte, die das Recht auf Funktionen wie den Standort oder das Adressbuch zuzugreifen, betreffen.
5. Profilbildung, das heißt Aspekte, die es ermöglichen aus den angegebenen und verarbeiteten Informationen ein Profil zu erstellen.
6. Werbung, das heißt Aspekte bei denen Werbung durch Drittanbieter ausgespielt oder aufgrund der Datenverarbeitung auf eine Person oder ein Gerät zugeschnitten wird.[49]
7. Übertragung an Dritte, das heißt Aspekte, die die Weitergabe von verarbeiteten Daten an Dritte betreffen.

Des Weiteren wurde in der Befragung unterschieden, ob der Datenschutzaspekt auf einer Webseite oder in einer App auftritt. Diese Unterscheidung wurde gemacht, um differenzieren zu können, ob bestimmte Datenschutzaspekte in Apps anders bewertet werden als auf Webseiten und möglicherweise eine andere Relevanz haben.

4.8.3.6.1.2 Ergebnisse: Überprüfung der Informationstexte-Cluster durch Online-Befragung

Die folgenden Abbildungen fassen die Bewertung der Teilnehmerinnen und Teilnehmer zusammen. Eine Abbildung zeigt jeweils die Bewertung der Schwierigkeit und hierbei den Anteil der Befragten, die finden, dass der Aspekt (Cluster) „schwierig“ oder „sehr schwierig“ zu erkennen bzw. verstehen ist; und eine Abbildung zeigt jeweils die Bewertung der Wichtigkeit und stellt dar, welcher Anteil der Befragten feststellt, dass es „wichtig“ oder „sehr wichtig“ ist, den Aspekt zu erkennen bzw. verstehen.

4.8.3.6.1.3 Schwierigkeit

Insgesamt lässt sich feststellen, dass die einzelnen Informationstexte-Cluster im Durchschnitt eher schwierig zu erkennen bzw. verstehen sind. Besonders die Cluster **„Übertragung an Dritte“** (Webseiten: 74% mind. „schwierig“; Apps: 74% mind. „schwierig“), **„Datenschutzerklärungen“** (Webseiten: 69% mind. „schwierig“; Apps: 72% mind. „schwierig“) und **„Profilbildung“** (Webseiten: 68% mind. „schwierig“; Apps: 69% mind. „schwierig“) werden von den Teilnehmerinnen und Teilnehmern als schwierig be-

[49] Der Aspekt „Werbung“ wurde in zwei Teilfragen erhoben. Zum einen wurden Teilnehmerinnen und Teilnehmer gebeten zu bewerten wie einfach und wichtig es für sie ist Werbung zu erkennen. Die zweite Teilfrage bezog sich auf personalisierte Werbung, das heißt Werbung, die auf persönlichen Daten basiert und an diese angepasst wurde.

wertet. Im Gegensatz dazu wird der Aspekt **„Werbung erkennen“** als weniger schwierig bewertet (Webseiten: 36% mind. „schwierig“; Apps: 40% mind. „schwierig“) und auch der Aspekt **„Zugriffe“** in Apps erkennen, wird im Gegensatz zu den erstgenannten Aspekten als weniger schwierig bewertet (58%).

4.8.3.6.1.4 Wichtigkeit

Auch bei der Wichtigkeit gibt es einen eindeutigen Trend. So bewerten die Teilnehmerinnen und Teilnehmer es im Durchschnitt eher wichtig, die Aspekte / Cluster zu erkennen bzw. verstehen. Die Bewertungen liegen bei fast allen Clustern bei ca. 80% mind. „wichtig“. Einzig das Erkennen von **„Werbung“** und **„personalisierter Werbung“** erhalten eine geringere Bewertung, jedoch werden auch diese Aspekte im Schnitt als wichtig bewertet.

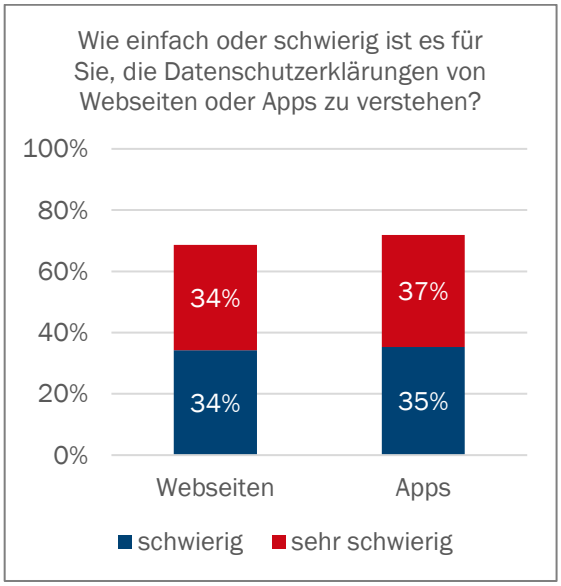


Abbildung 21: Schwierigkeit Cluster 1 „Datenschutzerklärungen“ (N=1.000)

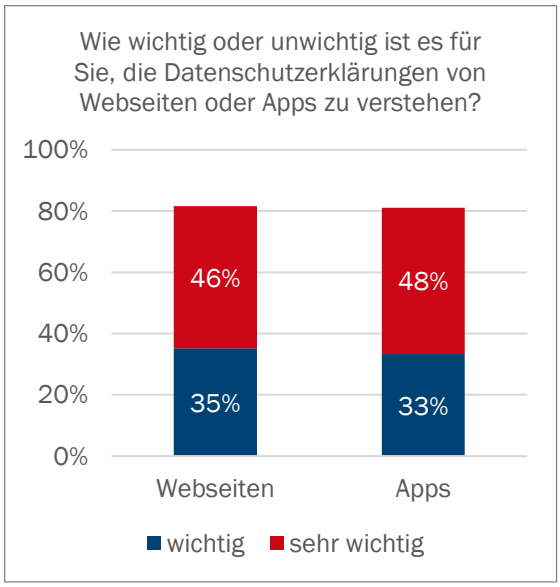


Abbildung 22: Wichtigkeit Cluster 1 „Datenschutzerklärungen“ (N=1.000)

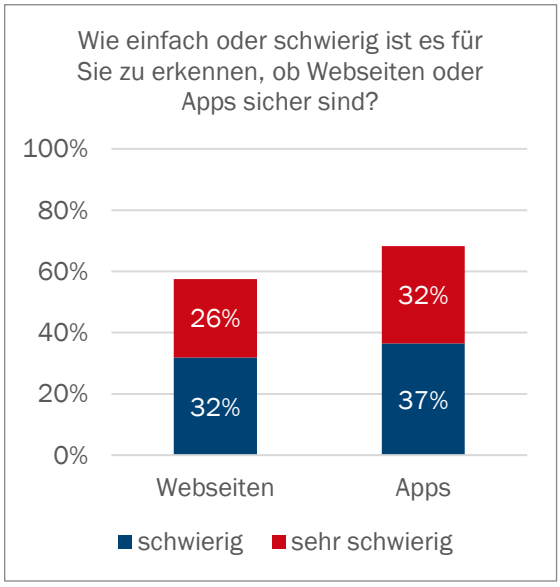


Abbildung 23: Schwierigkeit Cluster 2 „Sicherheit“ (N=1.000)

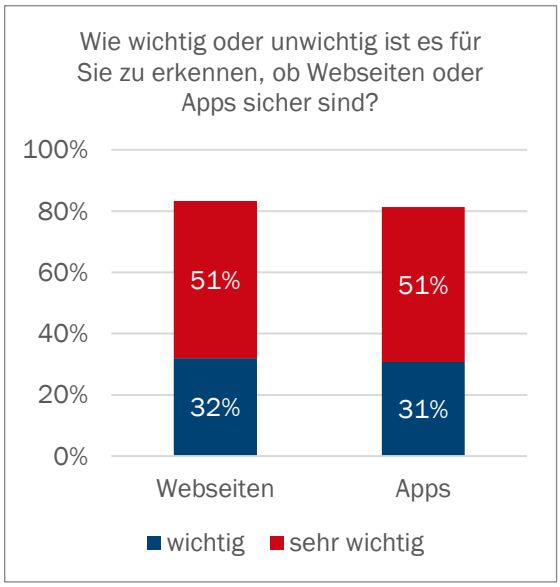


Abbildung 24: Wichtigkeit Cluster 2 „Sicherheit“ (N=1.000)

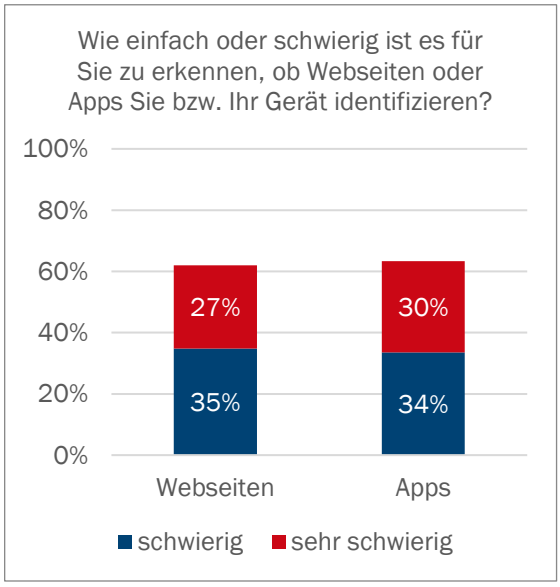


Abbildung 25: Schwierigkeit Cluster 3 „Identifikation“ (N=1.000)

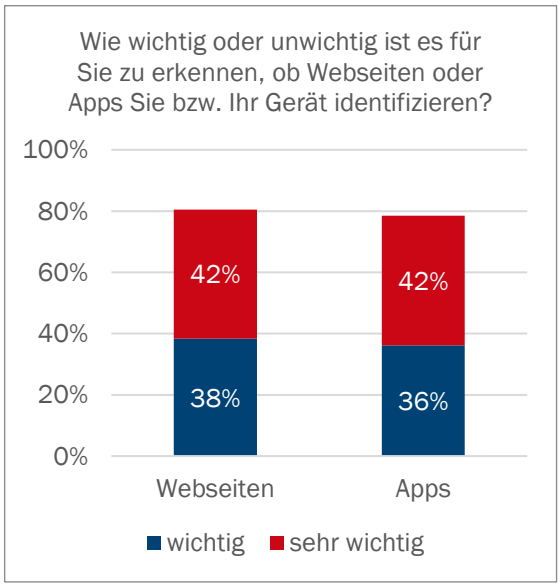


Abbildung 26: Wichtigkeit Cluster 3 „Identifikation“ (N=1.000)

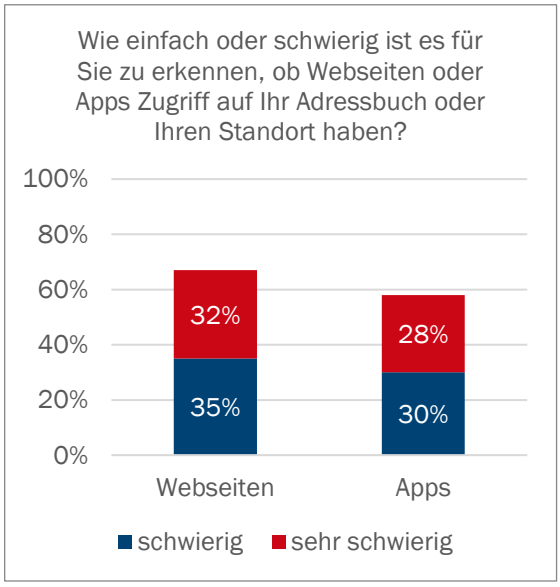


Abbildung 27: Schwierigkeit Cluster 4 „Zugriffe“ (N=1.000)

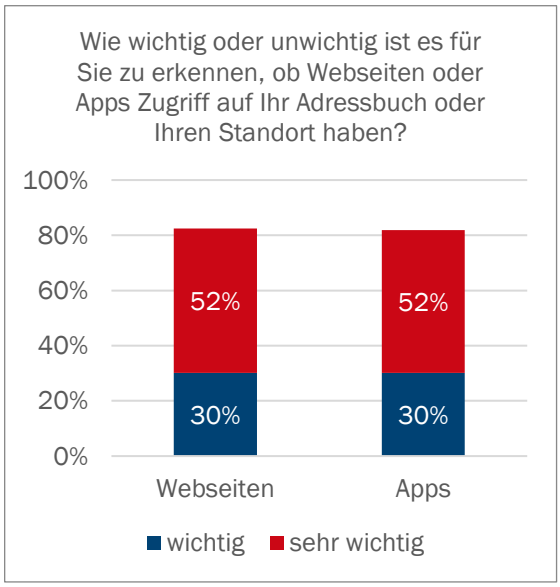


Abbildung 28: Wichtigkeit Cluster 4 „Zugriffe“ (N=1.000)

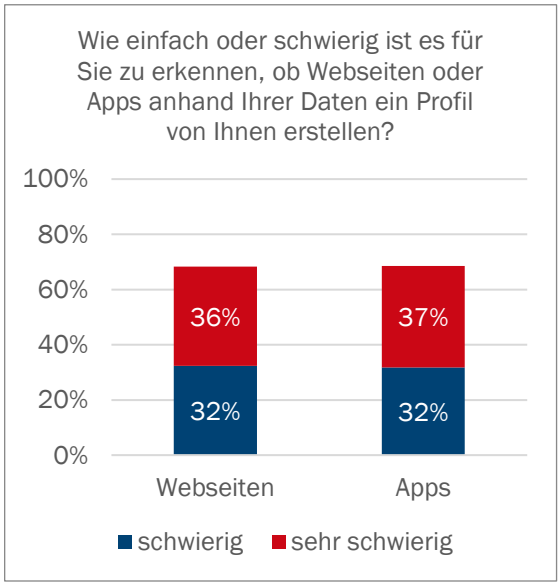


Abbildung 29: Schwierigkeit Cluster 5 „Profilbildung“ (N=1.000)

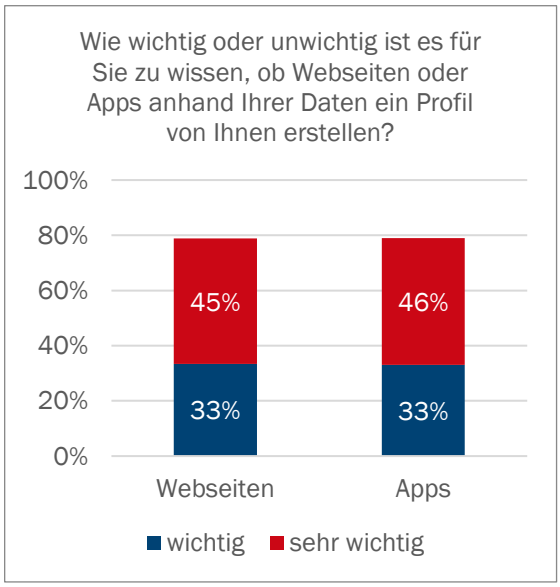


Abbildung 30: Wichtigkeit Cluster 5 „Profilbildung“ (N=1.000)

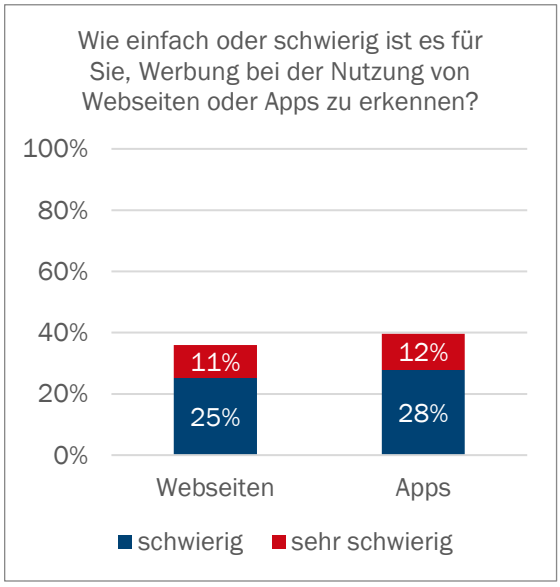


Abbildung 31: Schwierigkeit Cluster 6 „Werbung“ (1/2) (N=1.000)

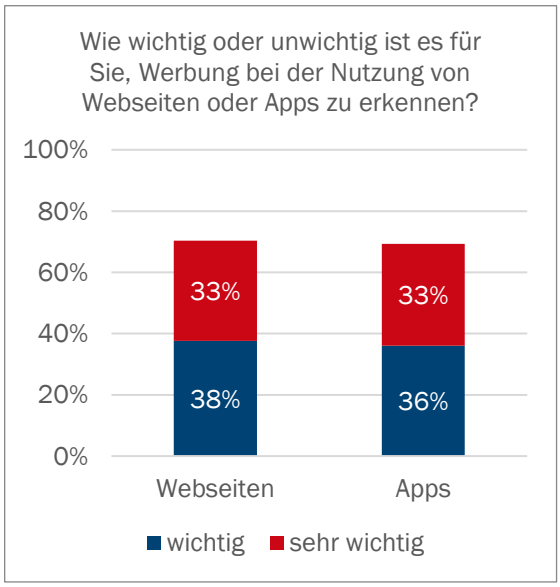


Abbildung 32: Wichtigkeit Cluster 6 „Werbung“ (1/2) (N=1.000)

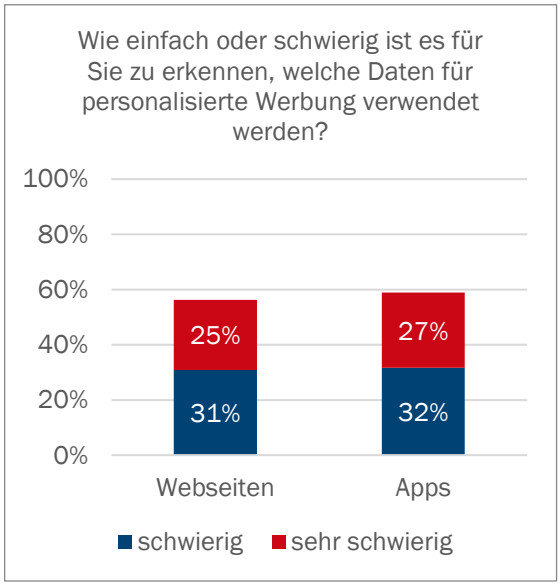


Abbildung 33: Schwierigkeit Cluster 6 „Werbung“ (2/2) (N=1.000)

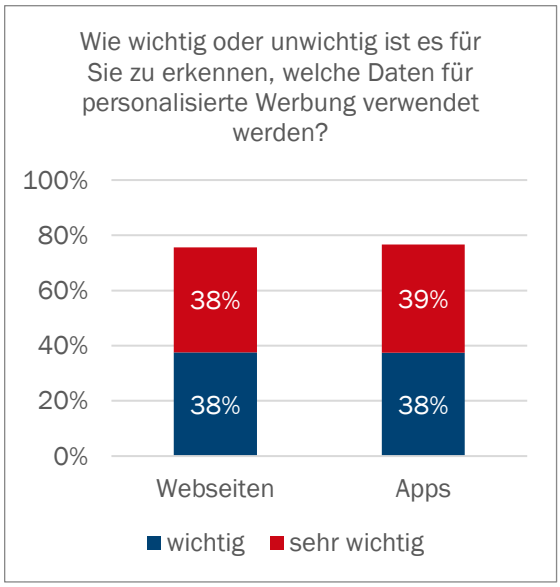


Abbildung 34: Wichtigkeit Cluster 6 „Werbung“ (2/2) (N=1.000)

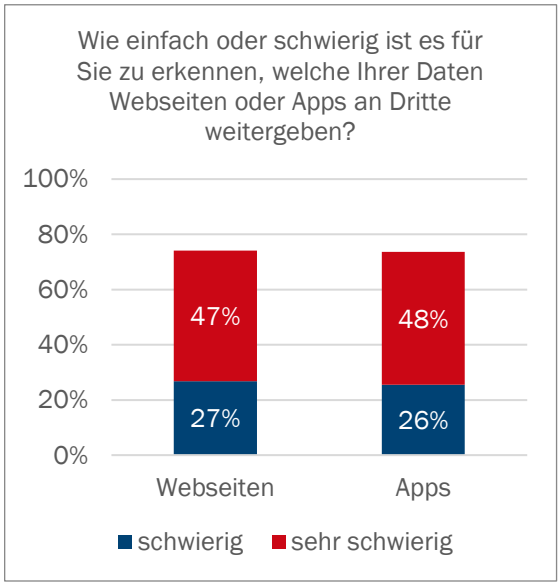


Abbildung 35: Schwierigkeit Cluster 6 „Übertragung an Dritte“ (N=1.000)

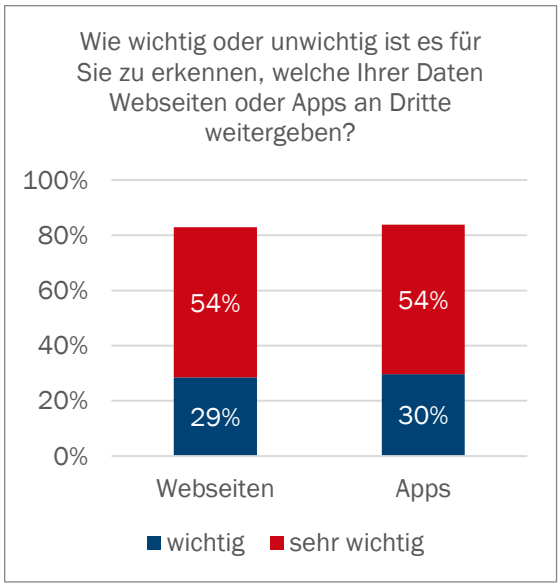


Abbildung 36: Wichtigkeit Cluster 6 „Übertragung an Dritte“ (N=1.000)

Die Ergebnisse werden durch die durchschnittlichen Bewertungen der einzelnen Cluster bestätigt. Diese finden sich in den Abbildung 37 und Abbildung 38.[50]

[50] Zur besseren Interpretierbarkeit wurden die Antwortkategorien metrisch gelabelt. Die Antwortkategorie „sehr schwierig“ entspricht auf der Normalisierungsskala einem Wert von -10, „schwierig“ entspricht einem Wert von -3 1/3, „einfach“ einem Wert von + 3 1/3 und „sehr einfach“ einem Wert von +10. Gleichermaßen wurden auch die Antwortkategorien zur Wichtigkeit der einzelnen Cluster belabelt. Sie reichen von „sehr unwichtig“ mit -10 bis „sehr wichtig“ mit +10. Somit können Mittelwerte größer 0 einer Bewertung von „eher einfach als schwierig“ bzw. „eher wichtig als unwichtig“ zugeordnet werden. Mittelwerte kleiner 0 korrespondieren mit einer durchschnittlichen Bewertung von „eher schwierig als einfach“ bzw. „eher unwichtig als wichtig“.

Den Abbildungen kann entnommen werden, dass im Durchschnitt bei den meisten Clustern das Erkennen bzw. Verstehen der Verbraucherinnen und Verbraucher als eher schwierig bewertet wird. Eine Ausnahme bildet das Cluster „Werbung“. Hier ist die durchschnittliche Schwierigkeitsbewertung größer 0 und somit empfinden die Teilnehmerinnen und Teilnehmer das Erkennen von Werbung eher einfach als schwierig. Bei der Bewertung der Wichtigkeit der einzelnen Aspekte findet man durchschnittliche Bewertungen größer 0. Somit kann festgestellt werden, dass das Erkennen bzw. Verstehen der einzelnen Aspekte von den Befragten eher wichtig als unwichtig bewertet wird.

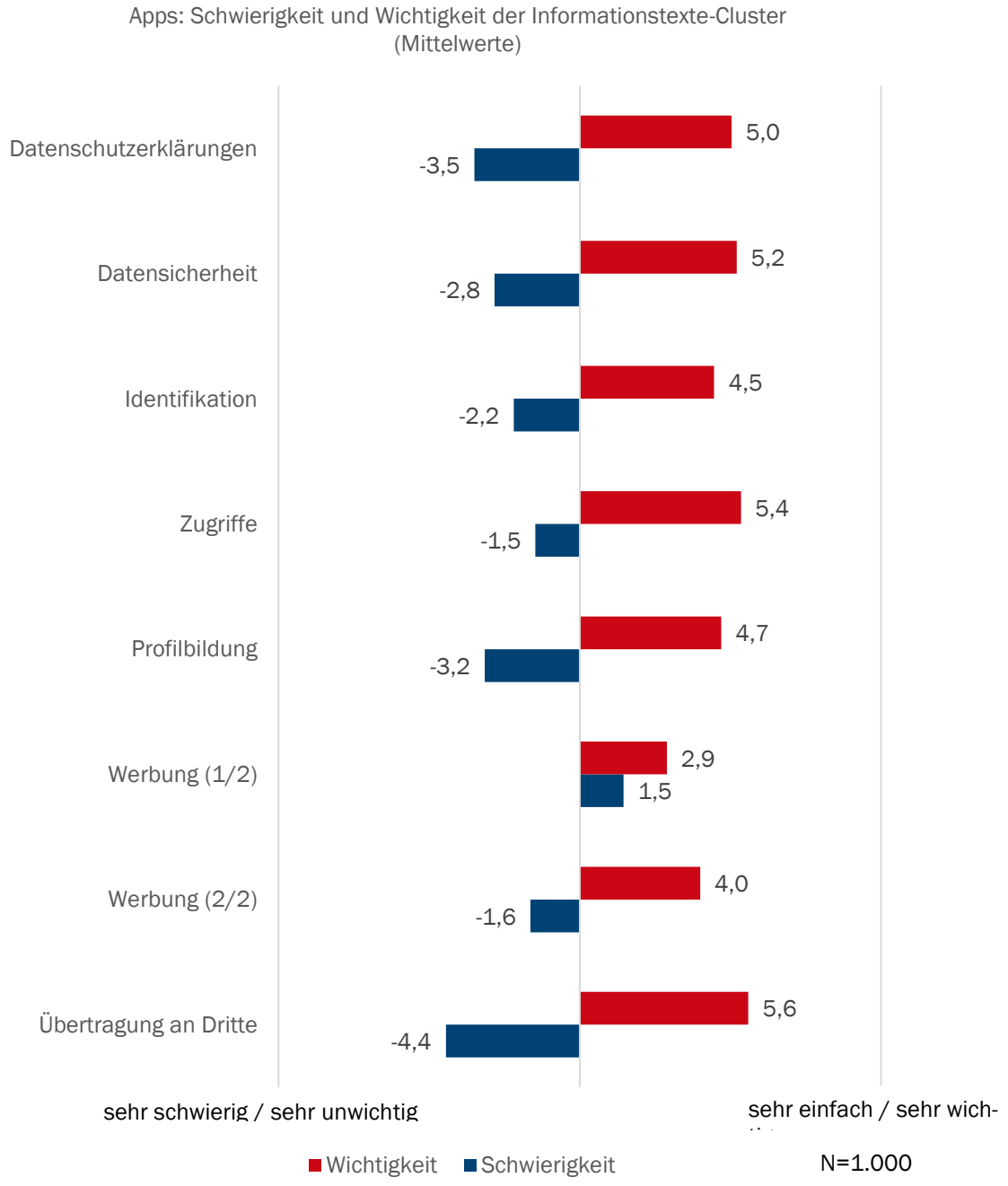


Abbildung 37: Schwierigkeit und Wichtigkeit der Informationstexte-Cluster (Mittelwert)

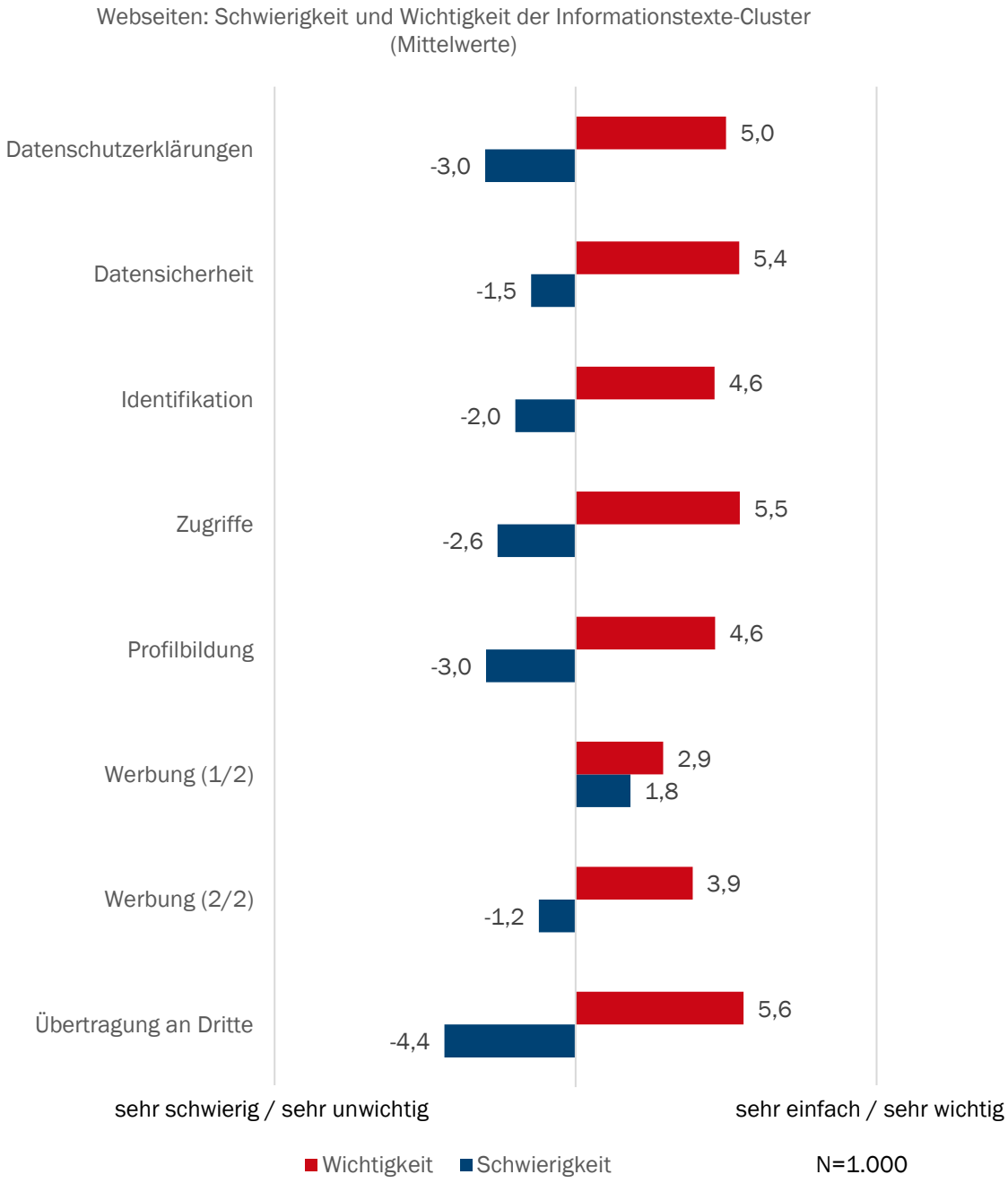


Abbildung 38: Webseiten: Schwierigkeit und Wichtigkeit der Informationstexte-Cluster (Mittelwert)

4.8.4 Zusammenfassung

Aus den Ergebnissen der Befragungen wurde als übergeordnete Zielsetzung (Mission und Vision) der Anwendung abgeleitet, dass die Labormuster des Projekts Nutzerinnen und Nutzern mehr Transparenz und Kontrolle beim Thema Datenschutz in ihren Apps liefern sollen. Um das Transparenzziel zu gewährleisten, wurden deshalb die Informationstexte erstellt, die die Ergebnisse der Befragungen und Fokusgruppengespräche berücksichtigen. Die Informationstexte folgen einem standardisierten Aufbau und

benennen Datenverarbeitungen konkret, erläutern Konsequenzen und beschreiben mögliche Handlungsoptionen, um die Kontrolle der Nutzerinnen und Nutzer zu steigern. Neben den Ergebnissen der technischen und semantischen Analyse (Kapitel 4.9 und 4.10) fließen sie als Kernbestandteil in die Labormuster (App-Client Kapitel 5 sowie Web-Anwendungen Kapitel 6) ein.

4.9 Technische Analyse der Datenverarbeitungen

Eine technische Analyse wurde im Forschungsprojekt für erforderlich erachtet, da eine ausschließliche Analyse der Datenschutzerklärungen als lückenhaft eingestuft wurde. Datenschutzerklärungen weisen – auch in rechtlich zulässigem Umfang – Unschärfen auf. Zudem ist es nicht ausgeschlossen, dass technische Änderungen nicht unmittelbar zu eigentlich notwendigen Anpassungen der Datenschutzerklärung führen.

Auch kommt es vor, dass Apps überhaupt keine oder widersprüchliche Datenschutzerklärungen bereitstellen. Die technische Analyse kann in diesen Fällen ergänzende als auch verifizierende Informationen zu den in der Datenschutzerklärung angegebenen Datenverarbeitungen bereitstellen.

4.9.1 Evaluierte Analysemethoden

Im Rahmen des Forschungsprojekts galt es auch die Datenverarbeitungen von Apps durch eine technische Analyse zu untersuchen. Hierbei stehen unterschiedliche Ansätze zur Verfügung, die getestet wurden und deren Realisierbarkeit im Rahmen des Projektes evaluiert wurden:

- manuelle detaillierte Analyse durch einen Penetrationstester
    - Nicht zielführend – Zeitaufwand nicht realistisch umsetzbar im B2C Markt
- manuelle Analyse des Netzwerkverkehrs
    - Teilweise zielführend – punktuell sinnvoll, aber nicht flächendeckend realisierbar
- automatische statische Analyse der Installationsdatei (Binary Analyse)
    - Zielführend – flächendeckende Informationsgenerierung möglich
- automatische dynamische Analyse der Installationsdatei
    - Zielführend – technische Hürden noch ohne Lösungsansatz

Für die technische Analyse wurde im Projekt auf bestehende Techniken zurückgegriffen, soweit dies dem Projektziel förderlich schien. So konnte insbesondere am Anfang auf die bereits bestehende und im B2B-Kontext eingesetzte Technologie APPVISORY™ AppScan Nucleus Engine zurückgegriffen werden.

Für die im Projektvorhaben benötigte Informationsgrundlage zu jeder einzelnen App wurde auf Basis der oben erwähnten Möglichkeiten eine Kaskade von unterschiedlichen Prüftechnologien für sinnvoll erachtet.

4.9.2 Eingesetzte Prüftechnologien

Die eingesetzten Prüftechnologien basieren dabei im Wesentlichen auf:

- Erfassung von Informationen aus offiziellen APIs (z.B. Apple-iTunes-API)
- Erfassung von Informationen aus inoffiziellen APIs (z.B. Google Play Store API)
    - Stabilisierung im Rahmen des Forschungsprojektes
- Erfassen der Installationsdateien zur weiteren Analyse (Installationsdatei-Downloader)
    - Optimierung im Rahmen des Forschungsprojektes
- Analyse der Installationsdateien (Entpacken, Entschlüsseln, Stringmatching & Codeanalyse)
- Analyse der Stringmatching-Ergebnisse (GEO-Location der Server & TLS Analyse)
    - Erweiterung der TLS-Analyse im Rahmen des Forschungsprojektes
- Analyse der Codeanalyse-Ergebnisse (Berechtigungen, Drittanbieter & Übertragung an Dritte)
    - Erweiterung um tatsächlich genutzte Berechtigungen, Überprivilegierung, Drittanbieter und Übertragung an Dritte
- Malwareanalyse (Vergleich mit bekannten Malware Fingerprints)

Die im Rahmen des Forschungsprojektes entwickelten Zusatzaspekte der nutzungsbasierten „Permission-Usage“ Attribute sind dabei eines der wichtigsten Elemente beim Zusammenführen der Informationen der semantischen und der technischen Analyse.

Bei der Malware Analyse wurde es für sinnvoller erachtet, auf bestehende Techniken und Datenbanken zurückzugreifen, da hierzu bereits hinreichend verlässliche Angebote auf dem Markt existieren. Für den Demonstrator wurde auf die bereits in der Nucleus Engine enthaltenen Informationen zurückgegriffen. Prinzipiell lässt sich dieses Modul aber durch jede andere Malware Analyse austauschen und kann daher als „Proof-of-Concept“ betrachtet werden.

4.9.3 Erweiterung der Prüftechnologien

Im Laufe des Projektes galt es weitere bestehende Techniken und deren Mehrwerte für das Projekt zu eruieren beziehungsweise die durch den Konsortialpartner mediaTest digital eingebrachte Analysetechnik für die Forschungsziele, wie oben aufgezählt, zu erweitern und anzupassen.

Insbesondere mussten hierbei neue Prüfparameter entwickelt werden, die einerseits die für Verbraucher spezifischen Informationen ermitteln und andererseits die aus der semantischen Analyse (siehe 4.10) stammenden Prüfergebnisse komplementieren beziehungsweise abgleichen konnten.

Diese, auf Verbraucherinnen und Verbraucher ausgerichtete Analyse, erforderte folgende Prüfkriterien zu überarbeiten:

4.9.3.1 iOS

- Ermittlung der für die Prüfung erforderlichen Informationen (App-Portfolio)

Es galt zu prüfen, ob und inwieweit die für die weitere Prüfung erforderlichen Daten in Form eines App-Portfolios ermittelt werden können. Da sich die Entwicklung auf einen Android-Client konzentrierte, wurden hinsichtlich iOS theoretische Überlegungen angestellt[51]. Welche Informationen in Rahmen eines solchen App-Portfolio notwendig waren und wie diese beispielsweise unter Android ermittelt werden konnten, wird unter 4.9.3.2 und 5.4 behandelt.

- Überwindung eingeschränkter Analysemöglichkeiten auf Grund einer Abhängigkeit von Jailbreaks

Eine Analyse der Installationsdateien kann nur erfolgen, wenn ein Jailbreak für iOS zur Verfügung steht. Nur so können analysierbare unverschlüsselte Installationsdateien erstellt werden; unter iOS werden alle Apps explizit für das Zielgerät verschlüsselt; eine universelle Entschlüsselungsmethode ist noch nicht bekannt. Um eine automatisierte und performante Abarbeitung der zu analysierenden iOS Apps realisieren zu können ist es auch notwendig, dass dieser Jailbreak für die von der App minimal geforderte iOS Version verfügbar ist. Hierbei wurde im Rahmen des Forschungsprojektes getestet, ob es möglich ist diese Einschränkungen zu entschärfen. Tatsächlich war es (manuell) möglich bei vielen z.B. iOS 12 voraussetzenden Applikationen, diese Voraussetzung auf eine niedrigere Version z.B. iOS 10 herabzusetzen. Dadurch ließe sich die Applikation zwar nicht stabil nutzen – es ließe sich aber ein Abbild der ausführbaren Datei erstellen, welches dann zur weiteren Analyse archiviert werden kann. Dieser Vorgang wurde auf Grund des sich im Forschungsvorhaben ergebenden Fokus auf Android nicht weiterverfolgt, aber als Konzept für eine mögliche Umsetzung und Automatisierung festgehalten.

4.9.3.2 Android

- Stabilisierung der inoffiziellen Google Play Store API

Mangels einer offiziellen Google Play Store API, müssen die inoffiziellen API Integrationen regelmäßig an (unangekündigte) Änderungen angepasst werden. Im Rahmen dieser Anpassungen wurden Stabilisierungen realisiert, die einen Abbruch der Analyse auch bei unerwarteten und fehlerhaften Antworten durch die API minimieren.

- Monitoring der inoffiziellen Google Play Store API

Zusätzlich zu den oben genannten Anpassungen wurde ein Monitoring in die Google Play Store API integriert, welches frühzeitig auf die Notwendigkeit etwaiger Anpassungen hinweist.

- Beschleunigung des Installationsdatei-Downloaders

Die im Forschungsprojekt zu entwickelnden Lösungen richten sich an Verbraucherinnen und Verbraucher. Verbraucherinnen und Verbraucher nutzen jedoch andere und in Summe mehr Apps– im Vergleich zu jenen, die im Unternehmensumfeld eingesetzt werden. So ist es in Letzterem eher unüblich, dass

[51] Siehe 5.3.2.2 sowie 5.1.

Spiele-Apps installiert werden. Unüblich ist es auch aufgrund der dort implementierten IT- und datenschutzrechtlichen Richtlinien, dass Apps ohne inhaltlichen Zusammenhang und ohne irgendwie geartete Freigabe installiert werden dürfen. Verbraucherinnen und Verbraucher indessen unterliegen solchen Beschränkungen nicht und installieren sich unter Umständen eine App als direkte Reaktion auf eine auf dem mobilen Endgerät ausgespielte Werbung. Somit ist mit deutlich diverseren Anfragen hinsichtlich der zu prüfenden Apps und deren jeweiligen App-Version zu rechnen. Mithin galt es den Downloader der entsprechenden Installationsdateien zu optimieren. Dieser wurde im Rahmen des Forschungsprojekt beispielsweise parallelisiert, um so eine schnellere Verarbeitung der Daten zu ermöglichen.

4.9.3.3 Generelle Analyse

- Erweiterung der TLS Analyse

Im Rahmen der TLS Analyse wurde auf eine offene, von Mozilla unterstützte TLS-Cipher-Suiten Bibliothek zur Bewertung der Qualität der eingesetzten Verschlüsselung gewechselt. Hierdurch wurde die Bewertung durch eine externe, verifizierbare und transparente Methode realisiert. Die Einführung neuer Verschlüsselungsmethoden – zum Beispiel TLS 1.3 – lässt sich durch die Aktualisierung der verwendeten Metadaten-Generierungs-Schicht integrieren ohne weiteren Aufwand zu verursachen. Die Server zur Durchführung dieser Analysen wurden optimiert, um eine automatische Durchführung dieser Tests zu beschleunigen. Hierdurch wurden Erkenntnisse, die außerhalb des Projekts gemacht wurden, auch für den DATENSCHUTZscanner verfügbar gemacht.

- Erweiterung der Berechtigungsanalyse

Auf Basis dieser technologischen Änderungen und Erweiterung der Analyseergebnisse galt es sinnvolle Ansätze zur Aus- und Bewertung dieser Ergebnisse zu evaluieren. Dabei gibt es unterschiedliche Varianten:

a. Einmaliges Auftreten einer Verarbeitung führt zur Abwertung (Verarbeitungsansatz)
b. Mehrmaliges Auftreten einer Verarbeitung führt zur Abwertung (Scoring Ansatz)
c. Tatsächlicher Abfluss führt zur Abwertung (Abfluss Ansatz)

Das einmalige Auftreten einer Verarbeitung ist ein zielführender Ansatz, um Datenverarbeitungen zu bestätigen bzw. zu negieren; dies eignet sich insbesondere im Forschungsprojekt, um Aussagen der semantischen Analyse zu bestätigen.

Das Gefahrenpotential von Applikationen auf Basis von wiederholten Datenzugriffen (Scoring Ansatz) zu verdeutlichen, bietet nur geringfügigen Mehrwert und wurde dementsprechend nicht implementiert.

Die Beurteilung der tatsächlichen Datenabflüsse ist der wertvollste Ansatz. Allerdings ist dieser nicht geeignet, die semantische Analyse in vollem Umfang zu bestätigen; einerseits können in einer Datenschutzerklärung lediglich potentielle Verarbeitungen genannt werden, andererseits können im Rahmen einer Datenflussanalyse nicht alle Datenabflüsse vollumfänglich, automatisch und zuverlässig ermittelt werden. Dies liegt unter anderem daran, dass Datenabflüsse nur unter bestimmten Nutzungsszenarien

oder nach bestimmten Zeitabläufen oder in Kombination mit bestimmten anderen, externen Einflüssen (zum Beispiel Ortswechsel des Endgerätes) ausgelöst werden können. Die Optionen solcher Auslöser sind so vielfältig, dass bereits eine Berechenbarkeit enorme Herausforderungen stellt.

Dementsprechend ist eine hybride Lösung für das Projektvorhaben verwendet worden. Das teilautomatisierte Prüfverfahren erzeugt mittels einer automatischen Auswertung sowie einer manuellen Prüfung an einem physikalischen Endgerät (Smartphone oder Tablet) einen umfangreichen Datensatz je App. Dabei umfasst das Prüfverfahren unter anderem die folgenden Aspekte:

- Automatische Erzeugung von App-Metadaten

Metadaten sind notwendig, um allgemeine Informationen über die geprüften Apps zum Beispiel im DATENSCHUTZscanner darzustellen und die installierte Version mit der geprüften zu verknüpfen. Eine solche Verknüpfung ist notwendig, um Betroffenen die zutreffenden Ergebnisse anzuzeigen, sollten sich diese zwischen den Versionen unterscheiden. Diese Informationen stammen im Wesentlichen aus den offiziellen Appstores des jeweiligen Betriebssystems.

- Automatische Datenstromanalyse

Hierbei werden alle in der App identifizierten möglichen Serververbindungen identifiziert und analysiert. Dazu werden die Standorte über eine GEO-Lokalisierung identifiziert. Der Serverbetreiber wird in erster Instanz aus dem Whois abgeleitet, wobei diese Information unzuverlässig sind (zum Beispiel bei Cloud Anbietern) und daher – soweit derartige Informationen vorhanden sind – als zweite Kaskade auf die Inhaber der Serverzertifikate zurückgegriffen wird. Zusätzlich werden bei verfügbarer Verschlüsselung der Serverkommunikation auch die Cipher-Suiten identifiziert. Dies lässt Rückschlüsse auf die Qualität der verwendeten Verschlüsselung zu. Eine Verbindung mit einem Server, der noch SSLv3 anbietet, wird beispielsweise als derart unzureichend verschlüsselt angesehen, dass dies sogar als unverschlüsselt interpretiert wird. Darüber hinaus wird innerhalb der Drittanbieterbibliothek, innerhalb der die möglichen Datenübertragungsziele identifiziert wurden, hinsichtlich einer angewandten Pseudonymisierung und Anonymisierung von Datensätzen geprüft. Hierbei wird unterschieden zwischen echter Verschlüsselung, Hashing und einfacher Konvertierung. Denn bei verschlüsselter Übertragung liegt am Zielort trotzdem das entschlüsselbare (unter Umständen personenbezogene) Datum vor, wobei beim Hashing nur der Vergleich mit bereits bekannten Daten möglich sein sollte.

Technisch ist eine Verschlüsselung somit darauf ausgelegt, Daten zunächst mit einem nur einem geschlossenen berechtigten Benutzerkreis bekannten Schlüssel vor dem Zugriff Unberechtigter zu schützen. Da den Berechtigten der Schlüssel bekannt ist, können die Daten von diesen auch stets zur weiteren Verarbeitung entschlüsselt werden.

Hashing zielt dagegen auf eine einmalige Veränderung der Daten. Eine Rückberechnung auf die Originaldaten ist aus einem Hashwert grundsätzlich nicht möglich, selbst wenn ein öffentlich bekannter Algorithmus zur Anwendung kommt. Jedoch ist hier zu unterscheiden, zwischen alten, sehr schwachen Algorithmen und jenen des aktuellen Stands der Technik. Auch bei alten Algorithmen ist zwar ein echtes „Zurückberechnen“ der Originaldaten nicht möglich, jedoch können alle erdenklichen Werte berechnet

und in einer Übersetzungstabelle (Rainbow Table) gespeichert werden. Ein Abgleich mit dieser Tabelle führt dann faktisch zu einer Möglichkeit, entsprechend schwache Hashwerte zu übersetzen.

Davon unabhängig sind Konvertierungen zu betrachten. Auch hierbei können Daten zwar für den menschlichen Leser auf den ersten Blick als verschlüsselt oder verhasht erscheinen. Technisch wurde aber keine der beiden Methoden angewendet. Angewendet wird z.B. „BASE64 encoding“. Konvertierungen stellen lediglich die Umwandlung in eine andere Struktur dar. Dies ist sinnvoll, wenn das gleiche Datum in unterschiedlichen Kontext verarbeitet wird und diese Kontexte alternative Aufbereitungen erfordern. Konvertierungen sind deshalb als Klartextübertragung zu bewerten.

Als Beispiel soll hier ein Zeitstempel dienen. Uhrzeiten werden – je nach Region – unterschiedlich dargestellt; auch haben etwaige Zeitzonen Einfluss auf die Art der Darstellung. Um zu vermeiden, dass ein Zeitstempel „30. Juni 2018, 12.00 Uhr (UTC)“ zu Verwirrung führt, wird dieser beispielsweise in einen UNIX Zeitstempel konvertiert. Der entsprechende UNIX Zeitstempel wäre: 1530360000.

- Punktuelle manuelle Datenstromanalyse

Bei auffälligen bzw. besonders wichtigen Datensätzen werden punktuell manuelle Analysen der tatsächlich aufgebauten Serververbindungen durchgeführt. Im Zuge dieser Datenstromanalyse können Aussagen zur Effektivität der Verwendung von „Certificate-Pinning“ und den tatsächlich aufgebauten Serververbindungen getätigt werden. Die ermittelten Ergebnisse werden nach dem Vier-Augen-Prinzip kontrolliert, verifiziert und sodann bei Bestätigung in die Prüfergebnisse übernommen.

- Berechtigungsanalyse

Ein erster Überblick der geforderten Berechtigungen kann schon den Metadaten entnommen werden, wobei sich hier nur die geforderten Zugriffs-Berechtigungs-Gruppen darstellen lassen, die nur einen groben Überblick über die eigentlichen Zugriffsberechtigungen der App ermöglichen. Nach der automatischen Analyse der Installationsdatei können weitere, detailliertere Aussagen über die tatsächlich verwendeten Berechtigungen getroffen werden (vgl. Beispiel unter 4.9.3.4). Im Rahmen der erwähnten punktuellen manuellen Analyse können zudem die während der Nutzung der geprüften App angeforderten Berechtigungen bezüglich ihrer Plausibilität für die Funktionalität der App bewertet werden.

4.9.3.4 Automatisierte Analyse von Installationsdateien

Zur effektiveren Vorhersage von möglichen Datenabflüssen wurde ein System zur Analyse der Installationsdateien realisiert, welches über die Berechtigungen hinaus auch die tatsächlich verwendeten Systemaufrufe identifiziert und dokumentiert. Dabei werden mögliche vom Programm verwendete Aufrufe festgehalten, um deren Verwendung zu identifizieren und dadurch die konkrete Kritikalität der genutzten Funktionen einschätzen und kommunizieren zu können.

Diese Analyse basiert auf unterschiedlichen Methoden. Jede dieser Methoden kann bereits sachdienliche Informationen ermitteln. In der Kombination können einzelne Prüfergebnisse zudem plausibilisiert werden, um eine hinreichende Verlässlichkeit der Prüfergebnisse sicherstellen zu können.

Einerseits wird anhand eines Stringmatchings nach bestimmten Schlagworten beziehungsweise Zeichenfolgen (inklusive aussagekräftiger URLs) gesucht, die Rückschlüsse auf spezifische Programmanweisungen zulassen. Aus diesen Programmanweisungen können entsprechende Datenverarbeitungen abgeleitet werden.

Eine weitere Möglichkeit besteht darin, implementierte Funktionsbibliotheken zu ermitteln. Hierbei kann auf die konkreten Bezeichnungen der Funktionsbibliotheken oder aber auch auf bestimmte, in der Installationsdatei – welche selbst eine Zusammenstellung vieler Dateien ist – enthaltene Dateinamen zurückgegriffen werden. Von den implementierten Funktionsbibliotheken kann ebenfalls auf potentielle Datenverarbeitungen geschlossen werden.

Eine weitere Methode besteht darin, bestimmte Konfigurationen oder Versionen häufig integrierter Bestandteile (zum Beispiel Funktionsbibliotheken) in einen Hashwert zu konvertieren. Ein Abgleich der Hashwerte mit entsprechenden Werten aus der Installationsdatei lässt sodann weitere Rückschlüsse auf die Datenverarbeitungen zu.

Beispiel:

Die bisherige Berechtigungsanalyse erlaubt eine Aussage bzgl. des Adressbuchzugriffs wie folgt:

> *Die Applikation fordert die Berechtigung „Adressbuch Zugriff“ inkl. lesendem und schreibendem Zugriff.*

Durch die Erweiterung der Analyse steht nicht mehr die Verfügbarkeit der Berechtigung im Fokus, sondern die Systemaufrufe, die z.B. *„ContactsContract.“* oder noch konkreter *„ContactsContract.CommonDataKinds.Phone“* beinhalten.

Danach sind durch die Identifikation dieser und weiterer Aufrufe auch Aussagen wie folgt möglich:

> *Die Applikation fordert die Berechtigung „Adressbuch Zugriff“ inkl. lesendem und schreibendem Zugriff.*
> *In der Applikation wird der lesende Zugriff auf die Telefonnummern im Adressbuch genutzt*
> *Schreibender Zugriff wurde nicht detektiert*

Diese detaillierteren Aussagen werden vollautomatisch realisiert und stehen ohne manuelle Analyse zur weiteren Verarbeitung zur Verfügung.

4.9.4 Ermittelbare Identifikationsmerkmale

Durch die Analyse der Installationsdateien konnte die Prüfung der Verwendung der nachstehenden Kriterien vollautomatisiert realisiert werden:

Kategorie	Prüfkriterium	Erklärung
Address Book Content	Adresse (Adressbuch)	Adressdaten eines Kontakts aus dem Adressbuch
	E-Mail-Adresse (Adressbuch)	E-Mail-Adresse eines Kontakts aus dem Adressbuch
	Name (Adressbuch)	Name eines Kontakts aus dem Adressbuch
	Telefonnummer (Adressbuch)	Telefonnummer eines Kontakts aus dem Adressbuch
Communication Data	E-Mail-Absender	Absender einer empfangenen E-Mail
	E-Mail-Betreff	Betreff einer gesendeten oder empfangenen E-Mail
	E-Mail-Inhalte	Inhalt einer gesendeten oder empfangenen E-Mail
	SMS-Absender	Telefonnummer des Absenders einer empfangenen SMS
	SMS-Inhalte	Inhalte gesendeter oder empfangener SMS
Device Account Data	Geräteaccount E-Mail-Adresse	E-Mail-Adresse des geräteeigenen Nutzeraccounts (z.B. Apple-Account, Google Play Account etc.)
	Geräteaccount Name	Zugehöriger Name des geräteeigenen Nutzeraccounts (z.B. Apple-Account, Google Play Account etc.)
	Geräteaccount Passwort	Zugehöriges Passwort des geräte-eigenen Nutzeraccounts (z.B. Apple-Account, Google Play Account etc.)
Device ID (dynamic)	Android Advertising ID	Werbe-ID zur Nutzeridentifikation, durch Nutzer änderbar. Nur gültig für Android-Geräte
	Gerätename	Vom Nutzer vergebener Name für das Endgerät
	IDFA (Identifier for Advertisers)	Werbe-ID zur Nutzeridentifikation, durch Nutzer änderbar. Nur gültig für iOS-Geräte
	IDFV (Identifier for Vendors)	ID zur App-/Nutzeridentifikation, vergeben nur für Apps des gleichen Herstellers. Nur gültig für iOS-Geräte

Device ID (static)	Android ID	Eindeutige Geräte-ID, durch Nutzer nicht änderbar. Nur gültig für Android-Geräte
	Geräte-Seriennummer	Eindeutige Seriennummer des Gerätes, durch Nutzer nicht änderbar
	GSFID (Google Service Framework ID)	Google-eigene ID zur Identifikation eines Gerätes im Google Play Store
	IMEI (International Mobile Equipment Identity)	Eindeutige Geräte-ID, durch Nutzer nicht änderbar
	UDID (Unique Device Identifier)	Eindeutige Gerätekennung, durch Nutzer nicht änderbar. Nur gültig für iOS-Geräte
Device ID (external)	WLAN MAC Adresse (verbundener Access Point)	Eindeutige Kennung des WLAN-Adapters eines verbundenen Access Points, durch Nutzer nicht änderbar
	WLAN SSID (verbundener Access Point)	Name oder Kennung des verbundenen Access Points
Device ID (public)	Bluetooth MAC Adresse	Eindeutige Kennung des Bluetooth-Adapters im Gerät, durch Nutzer nicht änderbar
	WLAN MAC Adresse (Gerät)	Eindeutige Kennung des WLAN-Adapters des Endgeräts, durch Nutzer nicht änderbar
	WLAN MAC Adresse (dummy)	Dummy aus Falschdaten für eine eindeutige Geräte-ID. Nur gültig für iOS-Geräte
Device Metadata	Betriebssystemversion	Betriebssystem und Versionsnummer des auf dem Endgerät installierten Betriebssystems
	Telefonmodell	Modellbezeichnung des genutzten Endgeräts
Device Password	Geräte-PIN	Ziffernfolge zum Entsperren des Endgeräts
	SIM-PIN	Entsperr-PIN der aktiven SIM-Karte
Device Password (external)	WLAN Passwort (verbundener Access Point)	Passwort des verbundenen Access Points
Exif Data	Foto-Metadaten	Metadaten auf dem Gerät gespeicherter oder übertragener Mediendateien

IP Address	Lokale IP Adresse (Gerät)	Lokale (interne) IP-Adresse des Endgeräts
	Öffentliche IP Adresse	Persönliche Adresse, zur Weitergabe bestimmt (z.B. an Logistik-Dienstleister)
Location Data	Standortdaten (exakt)	Exakte Standortdaten des Endgerätes, Genauigkeit auf wenige Meter
	Standortdaten (grob)	Grobe Standortdaten des Endgerätes, Genauigkeit ca. 1 km
Payment Data	BIC (Business Identifier Code)	Internationale Bank-Identifikationskennzahl
	IBAN (International Bank Account Number)	Internationale Kontonummer
	Kontoinhaber	Name des Kontoinhabers
	Kreditkarteninhaber	Name des Kreditkarteninhabers
	Kreditkartennummer	Kreditkartennummer
	PayPal E-Mail-Adresse	E-Mail-Adresse des PayPal-Accounts
	PayPal Passwort	Zugehöriges Passwort des PayPal-Accounts
Personal Data	E-Mail-Adresse	E-Mail-Adresse
	Name (Vor-/Nachname)	Vorname und/oder Nachname
	Persönliche Adresse	Persönliche Adresse, nicht öffentlich einsehbar (z.B. in einem Profil gespeichert)
	Telefonnummer	Telefonnummer
	Vertragsdaten (nicht sensitiv)	Nicht-sensible Daten ohne direkten Personenbezug (z.B. Buchungsnummern, Bestellnummern etc.)
	Vertragsdaten (sensitiv)	Sensible Daten mit direktem Personenbezug (z.B. Steuer-IDs, Versicherungsnummern etc.)
SIM ID (static)	Geräte-Telefonnummer	Telefonnummer der aktiven SIM-Karte
	ICCID (Integrated Circuit Card Identifier)	Seriennummer der aktiven SIM-Karte
	IMSI (International Mobile Subscriber Identity)	Eindeutige Nutzerkennung im Mobilfunknetz
Social Network Data	Facebook Login-Name	Login-Name des Facebook-Accounts
	Facebook Nutzername	Nutzername des Facebook-Accounts

	Facebook Passwort	Zugehöriges Passwort des Facebook-Accounts
	Google+ Nutzername	Nutzername des Google+-Accounts
	Google+ Passwort	Zugehöriges Passwort des Google+-Accounts
	Twitter Login-E-Mail-Adresse	Login-E-Mail-Adresse des Twitter-Accounts
	Twitter Nutzername	Nutzername des Twitter-Accounts
	Twitter Passwort	Zugehöriges Passwort des Twitter-Accounts
User Data	Benutzername	Benutzername
	Bundle Identifier (installierte App)	Eindeutiges Identifikationsmerkmal jeder App
	Passwort	Passwort
User Generated Data	Feedback und Kommentare	Durch Nutzer abgegebene Kommentare oder gesendetes Feedback
	Nutzer-Inhalte	Sammelkategorie für durch Nutzer erzeugte Daten (z.B. Notizen, gespeicherte Dateien, Texte etc.)
	Suchanfrage	Suchanfrage

Tabelle 4: Durch die technische Analyse ermittelbare Identifikationsmerkmale

4.9.5 Analyse integrierter Drittanbieter-Bibliotheken

In der Installationsdatei-Analyse wurde anlässlich des Forschungsprojekts auch die Erkennung von Drittanbieter-Bibliotheken erweitert. Es wird ein Unterschied bei den oben beschriebenen detektierten Zugriffen auf Daten zwischen der Hauptbibliothek und den integrierten Drittanbietern gemacht. Wird ein Zugriff gefunden, wird dieser dem *„Core“* oder einem Drittanbieter zugeordnet. Einzelne Drittanbieter werden dabei, basierend auf einer manuell erstellten Drittanbieter-Klassifikationsdatenbank, in die folgenden Kategorien einsortiert. Hierdurch wird die Zuordnung der integrierten Dienste und die mit diesen potentiell einhergehenden Datenverarbeitungen vereinfacht:

- Werbedienste: Einblendung von Werbeanzeigen innerhalb der App
- Analytics-Dienste: Messung von Nutzerverhalten, Performance-Analyse
- Tracking: Identifikation einzelner Nutzer/Geräte als Unterstützung zum Ausspielen von gezielter Werbung
- Cloud-Dienste: Online-Datenspeicher für Dateien auf unternehmensfremden Servern
- Kartendienste: Darstellung von Kartenmaterial, häufig zusätzlich mit Lokalisierung des Nutzers verbunden
- Soziale Netzwerke: Anbindung Sozialer Netzwerke wie Facebook, Xing etc. zum Teilen von Kommentaren, Dateien etc.

Diese Informationen lassen sich effektiv mit den Aussagen aus der Datenschutzerklärung abgleichen und dadurch die Konsistenz der in ebendieser gemachten Aussagen veri- bzw. falsifizieren.

4.9.6 Obfuskation

Es wurden verschiedene Verfahren analysiert, mögliche Obfuskationen, also das Verschleiern von Funktionen innerhalb des Quelltextes, zu umgehen. Insgesamt wurde hierbei kein zuverlässiger Ansatz identifiziert, der bei Integration zu einem Mehrwert innerhalb der Forschungsprojektes geführt hätte. Mithin können die durch die genutzte Technik ermittelten Ergebnisse absichtlich verschleierte Funktionen unter Umständen nicht ausweisen.

4.9.7 Zusammenfassung

Innerhalb des Forschungsprojektes konnten bestehende Analysemethoden modifiziert werden, sodass eine Vielzahl von relevanten Informationen über die Datenverarbeitungen einer App automatisiert ermittelt werden können. Hierbei wird zur Analyse auf die Installationsdatei zurückgegriffen. Obfuskationen sind zu befürchten; jedoch bestehen noch keine automatisierten und zielführenden Ansätze, einer solchen Obfuskation entgegenzuwirken.

Datenverarbeitungen können hierbei auf Basis von konkreten Programmanweisungen, integrierter Drittanbieter oder aber auch auf Basis bestimmter aufgerufenen URLs abgeleitet werden.

4.10 Semantische Analyse der Datenverarbeitungen

Eine Datenschutzerklärung sollte transparent und verständlich über die Datenverarbeitungen aufklären. Mithin sollten alle für Verbraucherinnen und Verbraucher relevanten Informationen in einer Datenschutzerklärung zu finden sein. Insbesondere gibt eine Datenschutzerklärung Aufschluss über die Verwendungszwecke und etwaige Übermittlungen und Rechtsgrundlagen für die jeweiligen Datenverarbeitungen. Diese Informationen sind durch eine rein technische Analyse gar nicht oder nur sehr bedingt zu ermitteln.

4.10.1 Überblick über einzelne Komponenten und deren Zusammenspiel

Die semantische Analyse umfasste mehrere Phasen. Diese können wie folgt eingeteilt werden:

- Auffinden von Datenschutzerklärungen, insbesondere in deutscher Sprache, von einer App oder einem Online-Service
- Datenschutzerklärung vorverarbeiten, indem störende Textstellen entfernt werden und der Text normalisiert wird
- Erstellen von Trainingsdaten durch Annotationen von ausgebildeten Juristen unter der Verwendung eines Active-Learning Tools.
- Training von Klassifikatoren wie zum Beispiel SVM's (SVM = Support Vector Machine)
- Anwenden der trainierten Klassifikatoren auf bisher unbekannte Texte für eine automatische Analyse
- Zusammenfassen der Annotationen zu für den Benutzer leicht verständlichen Informationstexten

- Speichern der Analyseergebnisse für den DATENSCHUTZscanner in einer Datenbank zur schnellen Abfrage
- Anzeigen der Analyseergebnisse in den Labormustern, also der DATENSCHUTZscanner App[52], dem PGuard Browser-Plugin[53] oder der Analyse-Webseite[54]

Folgende Grafik (siehe Abbildung 39) gibt diesen Ablauf noch einmal schematisch wieder und veranschaulicht das Zusammenspiel zwischen den verschiedenen Komponenten.

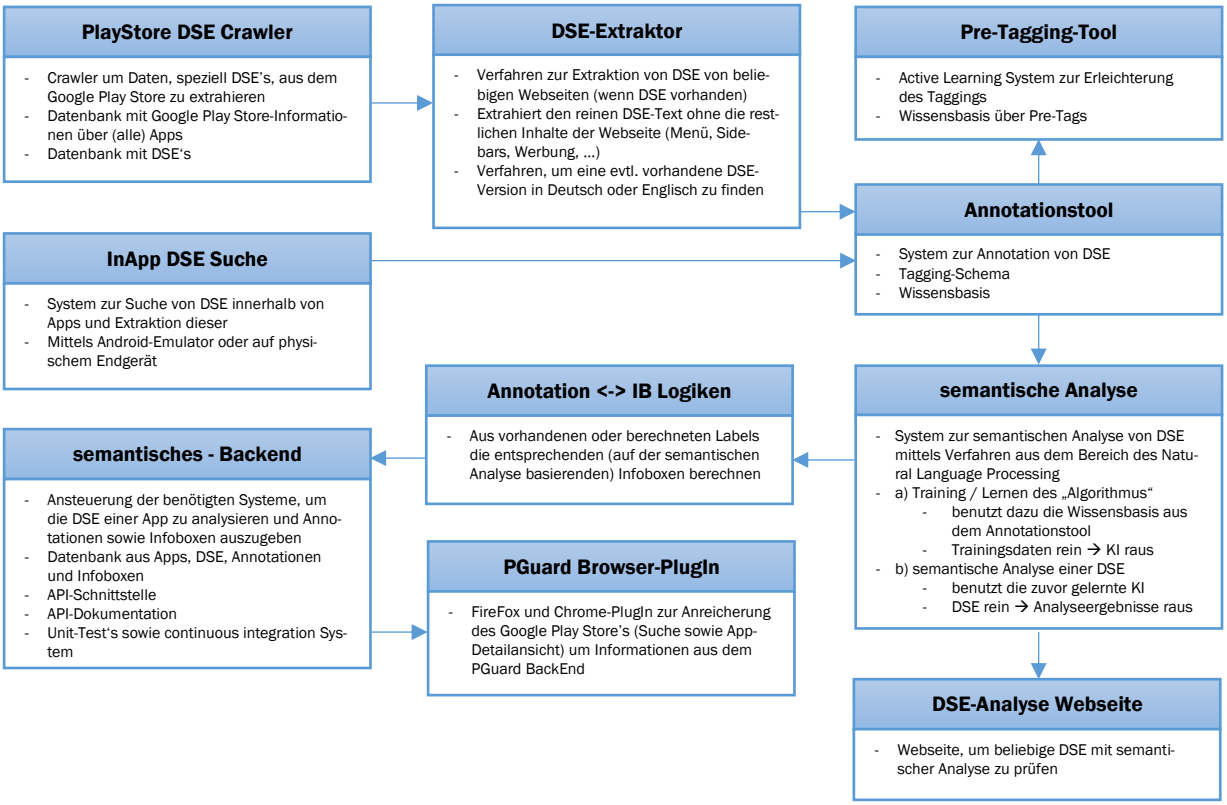


Abbildung 39: Schematischer Zusammenhang zwischen den einzelnen Komponenten der semantischen Analyse

In den folgenden Abschnitten wird jede Komponente noch einmal genauer erläutert. Anschließend erfolgt eine Evaluation der erreichten Qualität der semantischen Analyse. Hieran schließt ein Vergleich mit verwandten Forschungsarbeiten und -projekten. Abschließend wird ein Ausblick auf weitere Forschungsmöglichkeiten zur Verbesserung der Qualität gegeben.

[52] Siehe 5.
[53] Siehe 6.3.
[54] Siehe 6.2.

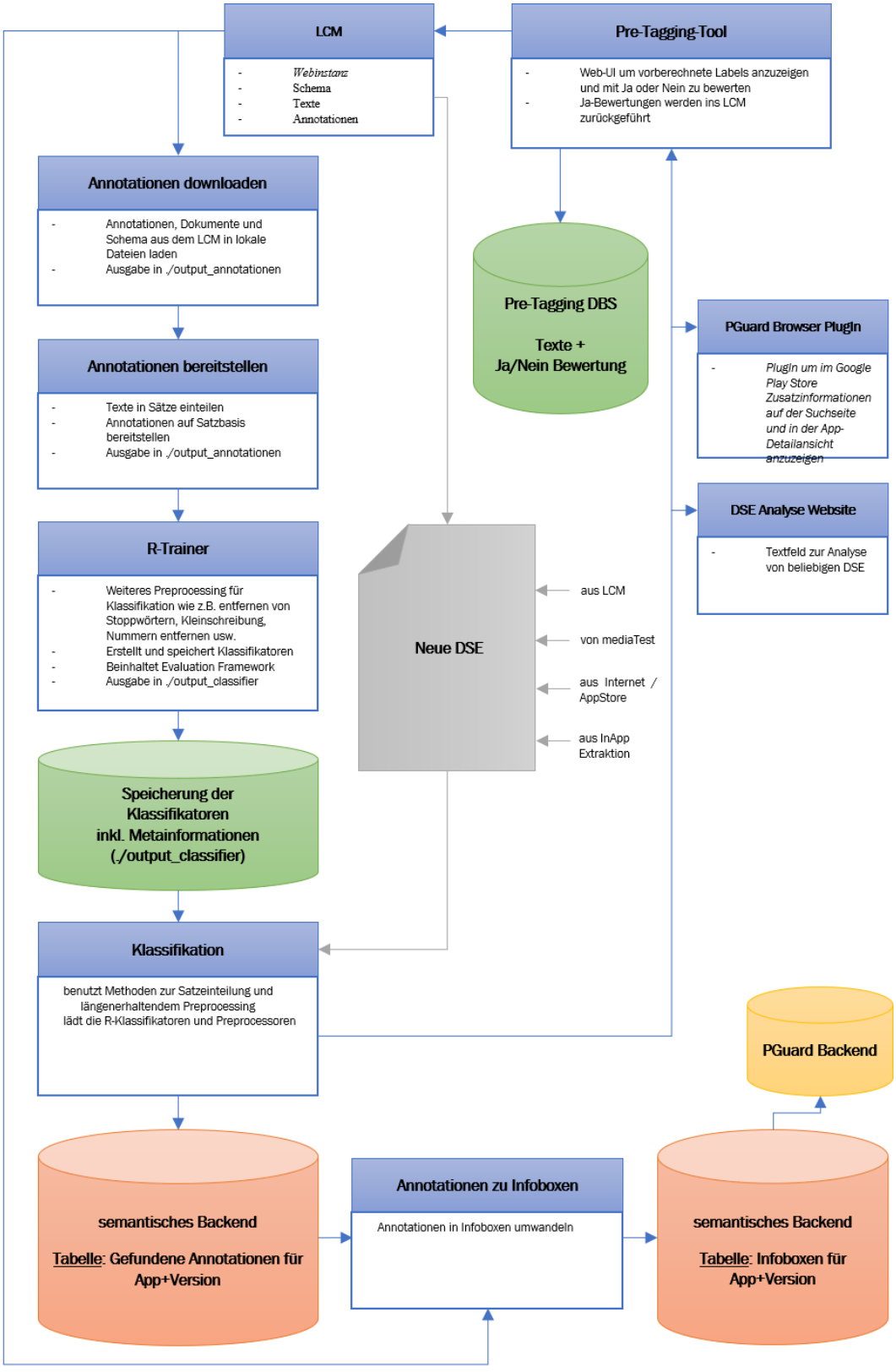


Abbildung 40: Schematischer Zusammenhang zwischen den einzelnen Komponenten der semantischen Analyse (alternative Darstellung)

4.10.2 PlayStore Datenschutzerklärungs-Crawler

Der PlayStore Datenschutzerklärungs-Crawler durchsucht den Google Play Store nach verlinkten Datenschutzerklärungen. Dabei startet der Crawler bei der Startseite des Google Play Store und besucht nach und nach alle verlinkten Unterseiten. Ist der Crawler aktuell auf einer Detailseite einer App, so extrahiert er alle Informationen zu dieser App und speichert sie in einer Datenbank. Besonders interessant sind hierbei folgende Informationen

- Bundle-ID der App
- Link zur Datenschutzerklärung, falls vorhanden (siehe auch Abbildung 41)
- Beschreibung der App
- Liste an ähnlichen Apps

Es wurden zum Stand Juni 2018 3,1 Millionen Apps gecrawlt. Von diesen 3,1 Millionen Apps verlinkten 1,2 Millionen Apps eine Datenschutzerklärung (ca. 37%). Von den 1,2 Millionen Links führten ca. 80.000 direkt zu einer deutschsprachigen Datenschutzerklärung.

Das Crawling des Google Play Store dauerte ca. 60 Tage. Durch den recht langen Abstand zwischen zwei Webseitenanfragen (ca. 1,7 Sekunden) wird nur wenig Last auf den Google Servern erzeugt und es ist daher nicht mit einer Sperrung des Crawlers zu rechnen.

Da der Crawler nicht Hauptfokus des Projekts war, wurde dieser nicht weiter optimiert.

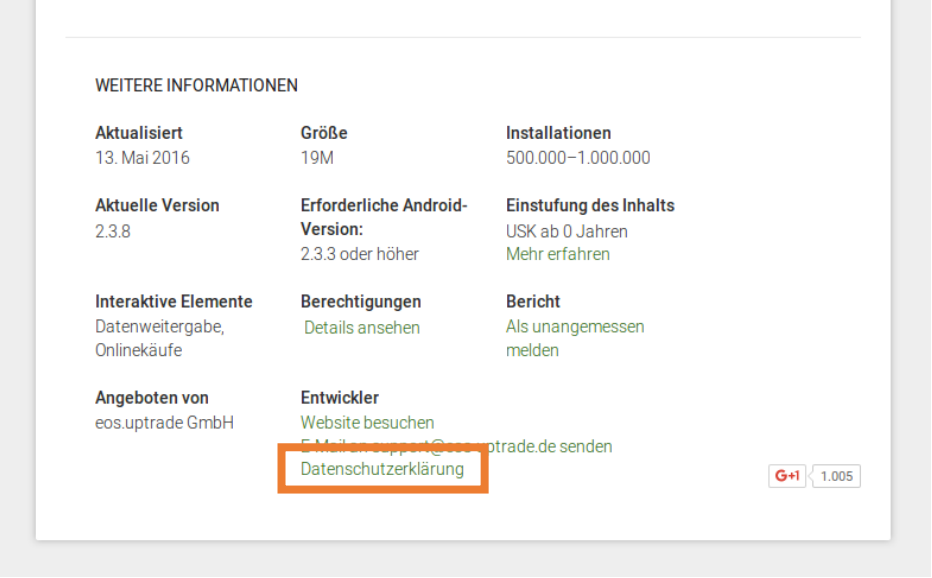


Abbildung 41: Verlinkung der Datenschutzerklärung für eine App im Google Play Store

4.10.3 Link-Guesser

Im Google Play Store wird oft eine englischsprachige Datenschutzerklärung verlinkt. Trotzdem kann eine deutschsprachige Version dieser Datenschutzerklärung vorliegen. Falls eine Datenschutzerklärung in

mehreren Sprachen vorliegt, gibt es meistens eine Auswahlbox (siehe Abbildung 42) für die Auswahl der bevorzugten Sprache. Mit Hilfe des „Link-Guesser“ sollte nun die URL zu der deutschsprachigen Version extrahiert bzw. erraten werden. Dabei wurden zwei Methoden angewendet.

4.10.3.1 Methode 1: Suche nach weiterführenden Links

Die erste Methode sucht im aktuellen Quelltext der Webseite nach Links zu weiteren Datenschutzerklärungen. Dabei werden alle Links (<a> und <iframe> HTML-Tags) und deren Ziele („href“ bzw. „src“ Attribut) nach bestimmten Schlüsselwörtern gefiltert:

- "privacy", "policy", "policies", "priva" , "eula", "terms", "/tos/", "/tou/"
- "datenschutz", "erklärung", "bedingung", "richtlinien", "agb"

Alle Links, die danach noch übrig sind, werden weiterhin regelbasiert bereinigt. Es werden dabei die Links aussortiert, die mit einer hohen Wahrscheinlichkeit nicht zu einer neuen Datenschutzerklärung führen. Dies kann zum Beispiel der Fall bei eingebundenen Social-Media-Share-Buttons auf einer Seite sein. Ausschlusskriterien können sich aus folgenden, in einer URL enthaltenen Texten ergeben:

- "share?url=", "twitter.com/intent/tweet?", "linkedin.com/shareArticle?", ...
- "/feed", "/login", "signin/", "/rss", "/answer/", ...
- "json", "pdf", "css", "js", "gif", "docx", "doc", "odt", "rtf", "zip", "rar", "bz2", "gz“, ...

Die restlichen Links führen mit einer hohen Wahrscheinlichkeit zu eine neuen, hoffentlich deutschsprachigen, Datenschutzerklärung.

4.10.3.2 Methode 2: Analyse und Modifikation der URL

Die zweite Methode sucht in der aktuellen URL nach Bestandteilen, die auf eine bestimmte Sprache der Webseite hinweisen. Dies ist oft ein „en“ für eine englischsprachige Webseite und „de“ für eine Deutschsprachige. Die folgende Tabelle gibt die gefundenen Schemata wieder:

Regexp	Beispiel
/(.*\/)(en|de)($)/	http://about.king.com/consumer-terms/terms/en
/(.*\/)(en|de)(\/.*)/	http://product.gree.net/us/en/privacy
/(https?\:\/\/)(www\.)?(en|de)(\..*)/	http://en.babybus.com/index/privacyPolicy.shtml
/(.*_)(en|de)(.*)/	http://www.notyx.com/web/policy/sopa_policy_en.xml
/(.*)(en-us |de-de)(.*)/	https://wwwsecure.lego.com/en-us/legal/legal-notice/privacy-policy
english → german	http://www.pg.com/privacy/english/privacy_notice.shtml

Tabelle 5: Schemata zur Analyse und Modifikation von URLs zum Auffinden deutscher Datenschutzerklärungen

In der aktuellen URL wird entsprechend des regulären Ausdrucks „en“, „en-us“ bzw „english“ durch die deutsche Entsprechung „de“, „de-de“ bzw. „german“ ersetzt. Anschließend wird geprüft, ob die daraus resultierende URL online verfügbar ist.

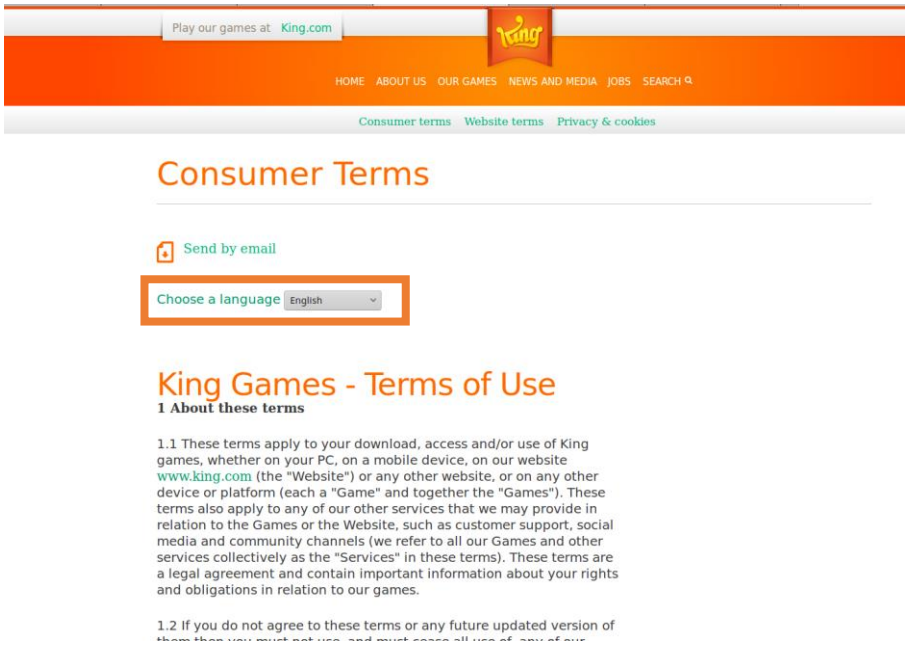


Abbildung 42: Beispiel für eine Datenschutzerklärung die in mehreren Sprachen verfügbar ist

4.10.4 Datenschutzerklärungs-Text Extraktor

Wurde eine Webseite mit einer Datenschutzerklärung gefunden, so muss die Datenschutzerklärung ohne störenden anderen Text extrahiert werden. Dies ist mitunter schwierig, da die Datenschutzerklärungen auf den Webseiten der einzelnen Hersteller und somit in einer stark heterogenen Struktur bereitgestellt werden (siehe Abbildung 42 und Abbildung 43).

Um die Datenschutzerklärung zu extrahieren wurden alle störenden Inhalte der Webseite, wie zum Beispiel Menüs, Sidebars, Navbars, News, Werbung usw. regelbasiert entfernt, sodass nur noch der Text der Datenschutzerklärung übrig bleibt. Dabei werden folgende Regeln angewendet (Auszug):

- Entferne zuerst folgende HTML-Elemente komplett:
  - "script", "noscript", "link", "comment()", "form", "head", "header", "footer", "foot", "nav", "style", "img", "input", "label", "select", "meta", "dl", "title", "button“, ...
- Entferne HTML-Elemente welche als Text folgenden Inhalt haben:
  - „zurück“, "back to top", "view full policy", "click here", "email protected", "viagra“, ...
- Entferne anschließend alle HTML-Tags die folgende id oder class haben:
  - "header", "header_banner", "impressum", "bottombar", "side_menu", "main_menu", "sub_menu", "social", "mailinglist", "advert", "ads", "promotion", "authorbox", "recentpost", „hidden“, ...
- Entferne die Elemente aber nicht, falls die id oder class folgenden Werte enthalten:
  - "with", "has", "not", "no" <-> "sidebar", "side_bar", "side-bar", "menu", ...

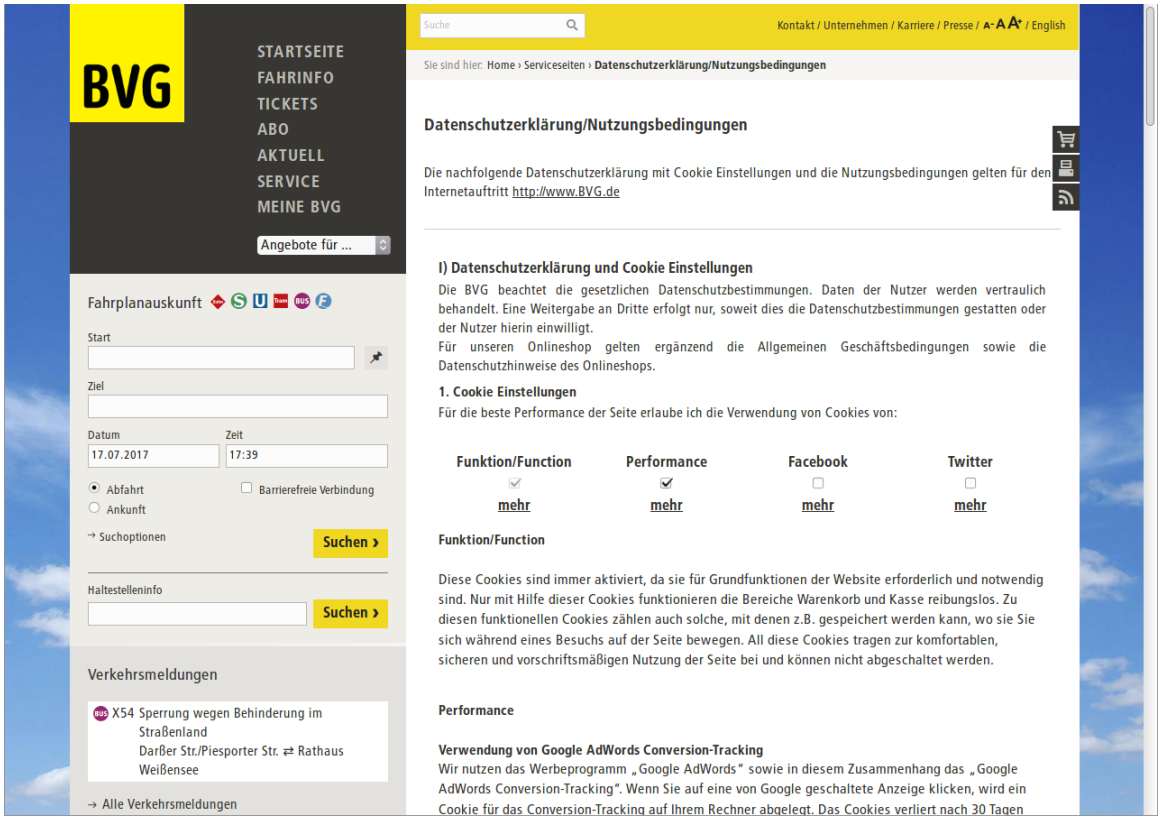


Abbildung 43: Beispiel für eine Webseite mit einer Datenschutzerklärung und vielen weiteren Inhalten.

4.10.5 InApp Datenschutzerklärungs-Searcher

Nicht für jede App ist eine Datenschutzerklärung im Google Play Store verlinkt. Eine Möglichkeit dennoch an eine Datenschutzerklärung zu ermitteln und auszuwerten, ist das Durchsuchen der App. Mit Hilfe des „InApp Datenschutzerklärungs-Searcher" wird eine App automatisch nach dieser durchsucht. Dabei wird die App in einem Emulator oder in einem per USB-Debugging angeschlossenen physischem Gerät installiert. Anschließend wird mit Hilfe des „UIAutomator" die App ferngesteuert und „durchgeklickt" (siehe Abbildung 44 und Abbildung 45). Wird eine Datenschutzerklärung gefunden, so wird diese zurückgegeben und der Suchvorgang ist beendet.

Das Auffinden einer Datenschutzerklärung innerhalb einer App ermöglicht auch einen Vergleich dieser integrierten Datenschutzerklärung mit der verlinkten Version im Google Play Store (falls vorhanden) und einen Hinweis für den Benutzer, falls Unterschiede festgestellt werden.

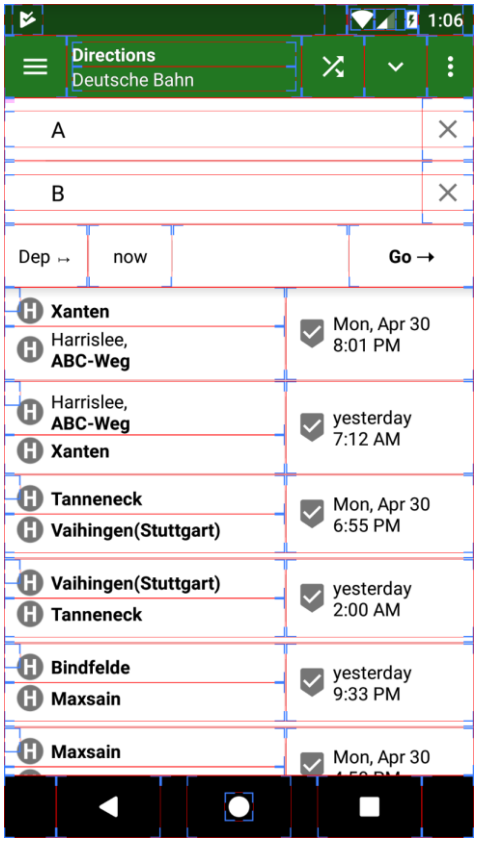


Abbildung 44: Beispielhafte Erkennung der UI-Elemente in einer Android App

```
<?xml version="1.0" encoding="utf-8"?>
<LinearLayout xmlns:android=
  "http://schemas.android.com/apk/res/android"
              android:layout_width="fill_parent"
              android:layout_height="fill_parent"
              android:orientation="vertical" >
    <TextView android:id="@+id/text"
              android:layout_width="wrap_content"
              android:layout_height="wrap_content"
              android:text="I am a TextView" />
    <Button android:id="@+id/button"
            android:layout_width="wrap_content"
            android:layout_height="wrap_content"
            android:text="I am a Button" />
</LinearLayout>
```

Abbildung 45: Beispielhafter Aufbau einer App im XML-Format

4.10.6 Tagging Tool

Für die semantische Analyse ist ein qualitativ hochwertiger Trainingsdatensatz erforderlich. Dieser wird mit Hilfe sogenannter Annotationen generiert. Alle Annotationen (siehe Abbildung 47) dienten später als Trainingsdaten für das Erstellen der Klassifikatoren.

Im Rahmen des Forschungsprojektes wurde unter anderem hinsichtlich der Annotationstools auf bestehende technische Lösungen für diesen Prozess zurückgegriffen. Die Wahl fiel hierbei auf den „Leipzig Corpus Miner“[55] (LCM), von der Abteilung für Automatische Sprachverarbeitung der Universität Leipzig,

[55] http://lcm.informatik.uni-leipzig.de/

mit dessen Hilfe Texte mit Annotationen versehen werden (siehe Abbildung 46) können. Der Rückgriff auf bestehende Technik ermöglichte es, sich im Projekt auf die für das Projekt relevanten Aspekte zu konzentrieren und für grundlegende Textanalysen ohne besonderen Fachkontext auf die Daten, Erkenntnisse und Entwicklungen dieser bestehenden Technik zurückzugreifen.

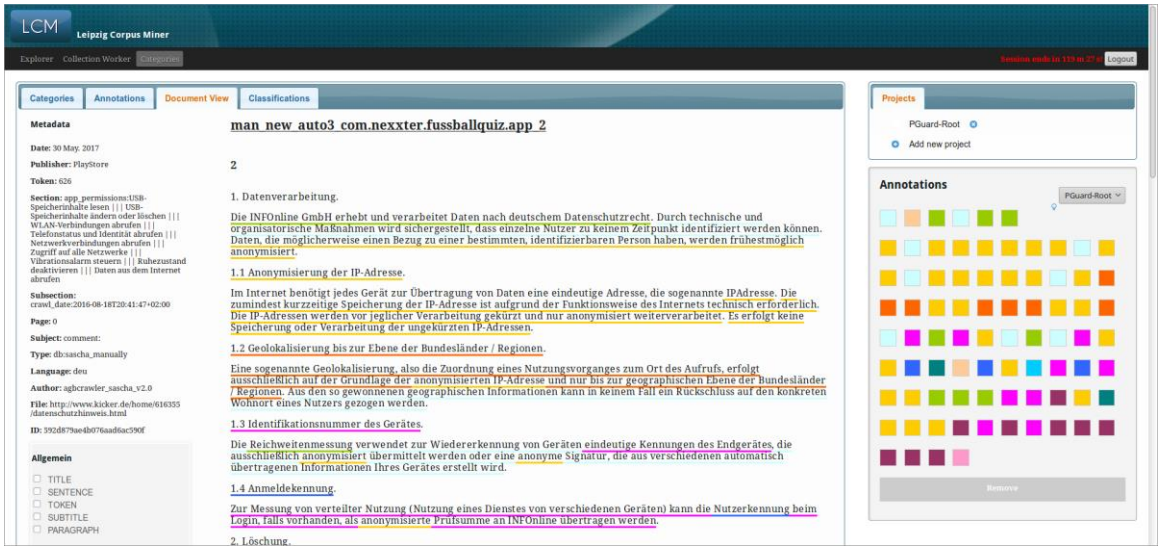


Abbildung 46: Beispielhafte Annotation einer Datenschutzerklärung im LCM.

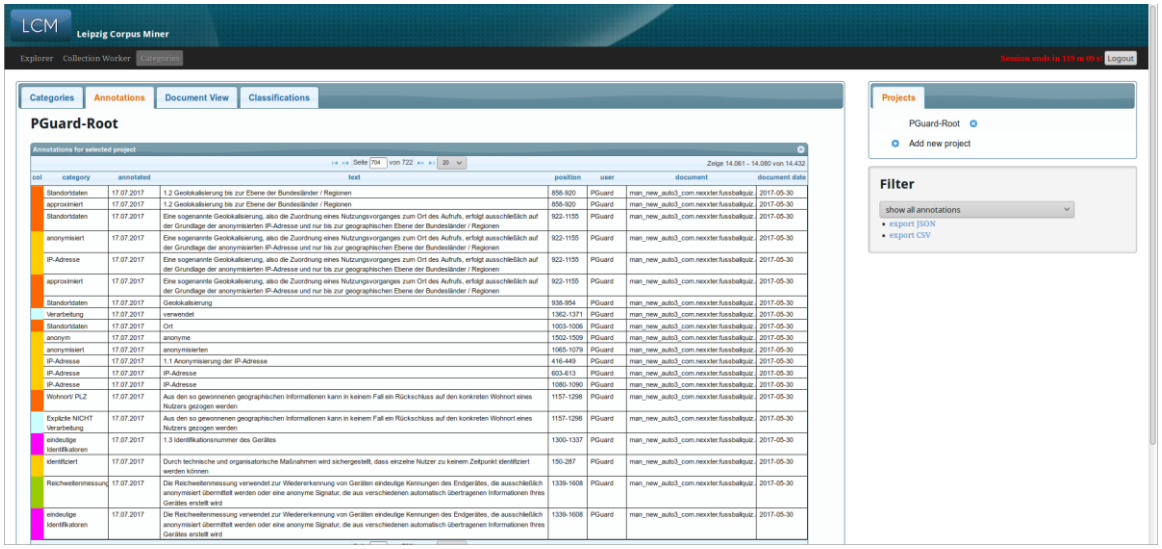


Abbildung 47: Ausschnitt über alle je gemachten Annotationen.

Für das Erstellen von Annotationen ist ein Annotationsschema nötig. Das in diesem Projekt verwendete Annotationsschema umfasst ca. 300 Labels.

4.10.6.1 Das Annotationsschema – Hintergrund

Das Annotationsschema wurde in enger Zusammenarbeit zwischen den im Projekt beteiligten Juristen und Informatikern entwickelt. Diese intensive Zusammenarbeit erschien notwendig, da die Qualität des Annotationsschemas und der daraus abgeleiteten Annotationen die Wertigkeit der im Projekt zu entwickelnden Algorithmen stark beeinflusst. Annotationen können unabhängig voneinander oder in Abhängigkeit zueinander verstanden werden. Etwaige Abhängigkeiten können durch nachgelagerte Logiken in den Algorithmen oder aber auch durch Kategorisierungen innerhalb des Annotationsschemas abgebildet werden.

Zunächst wurden die Ziele des Annotationsschemas erörtert. Hierbei sind die technischen Möglichkeiten und die technischen Erforderlichkeiten berücksichtigt worden. Im Forschungsprojekt galt es hierbei auch die unterschiedlichen, weiteren Disziplinen und deren Beiträge abzubilden. Eine voll-kontextualisierte Analyse (welche konkreten Daten für welche konkreten Zwecke) der jeweiligen Datenverarbeitungen wurde als sinnvoll, technisch aber anspruchsvoll eingestuft. Zugleich wurde der Mehrwert für Nutzerinnen und Nutzer aber bereits darin erkannt, dass diese bereits Informationen darüber erhalten, welche Daten verarbeitet werden und zu welchen Zwecken Daten verarbeitet werden – also ausdrücklich auch ohne automatisierte Verbindung dieser beiden Aussagen.

Aus technischer Sicht ist das „Ob überhaupt“ eine Information vorliegt (welche Daten, welche Zwecke) notwendige Voraussetzung um diese Informationen in möglichen Folgeverarbeitungen zu kontextualisieren. Etwaige Kontextualisierungen wurden zudem technisch nicht im unmittelbaren Bereich der Annotation, sondern eher im Bereich der nachgelagerten Folgelogiken verortet.

Vor diesem Hintergrund konnte sich das Annotationsschema auf die reine, inhaltliche Aufbereitung der Informationen konzentrieren und vorerst komplexe Abhängigkeiten vernachlässigen. Entsprechend der Informationstexte galt es allerdings das Annotationsschema stets veränderten Bedingungen anzupassen und neue Erkenntnisse und Herausforderungen zu adressieren. Hierbei war auch zu berücksichtigen, dass die semantische Analyse die für die Nutzerinnen und Nutzer aufbereiteten Informationstexte mit sachdienlicher Datengrundlage versorgen muss. Veränderungen im Rahmen der Informationstexte[56] führten somit zu einer Überprüfung der Notwendigkeit etwaiger Anpassungen im Annotationsschema.

4.10.6.2 Annotationsschema – Prozess der Entwicklung und kontinuierlichen Anpassung

Das Annotationsschema basierte anfänglich sehr stark auf juristischen Ansätzen, datenschutzrechtliche Aspekte zu gliedern und zusammenzufassen. Datenarten wurden in die – im juristischen auch im IT-sicherheitstechnischen Kontext – typischen Datenarten gegliedert und entsprechende Daten hierin gesammelt. Hierunter fallen zum Beispiel Allgemeine Persönliche Daten, besondere Arten von personenbezogenen Daten, Abrechnungs- und Bankdaten, Standortdaten, Kommunikationsdaten, Social Media Daten (User Generated Content).

[56] Siehe hierzu 4.8.3.2 bis 4.8.3.6.

Ergänzend galt es Verarbeitungszwecke zu erfassen. Auch hier wurde zunächst auf die im juristischen Kontext üblichen Kategorisierungen zurückgegriffen, da davon auszugehen war, dass Datenschutzerklärungen als eine besondere Form der Rechtstexte höchstwahrscheinlich Rechtssprache und somit die dortigen Typisierungen verwenden werden. Dies umfasst zum Beispiel „Vertragserfüllung“, „Analyse“, „Werbung“ sowie „Profilbildung“.

Letztlich galt es zudem den spezifischen Kontext von Apps und mobilen Endgeräten abzubilden. Mithin wurden Kategorien erstellt, die die Funktionsgruppenzugriffe, lokal gespeicherte Informationen wie zum Beispiel Kontaktdaten oder Bild- und Videodaten, oder aber auch Geräteinformationen abbilden.

Im Laufe des Projekts wurden diese Informationen um Kategorien und Annotationen ergänzt, soweit dies aufgrund der manuellen Analyse der Datenschutzerklärungen sinnvoll erschien. Dies umfasst insbesondere Ergänzungen um „Fitnessdaten“, „Gesundheitsdaten“ oder „Reisedaten“.

Gleiches erfolgte, wenn für die weitergehende, technische Verarbeitung der Annotationen ergänzende Informationen erforderlich schienen. Dies war im Besonderen der Fall bei Verneinungen oder „Wenn-Dann“-Phrasen.

Im Übrigen wurde das Annotationsschema häufig im Wege einer Konsolidierung überarbeitet. Hiermit rückte die Kategorisierung stückweise von der klassischen juristischen Kategorisierung ab und entwickelte seine eigene Logik, in der bestimmte Redundanzen durch Einführung eigener Meta-Informationen und Meta-Kategorien reduziert wurden.

4.10.6.3 Annotationsschema – Beispielhafte Abläufe des Annotierens

Es wurde immer das ausschlaggebende Wort bzw. die ausschlaggebende Phrase in einem Satz annotiert. Um Negationen abzubilden, wurde die betroffene Textstelle zunächst mit dem eigentlichen (positiven) Label (also bspw. „speichern“ mit „Verarbeitung“) markiert und danach dieselbe Textstelle noch mit „Explizite Nicht-Verarbeitung“ überlagert. Dies galt nicht nur für bestimmte Datentypen wie Standortdaten, sondern auch für Kategorien wie „Weitergaben an Dritte“. Erfolgt also zum Beispiel keine Weitergabe an Dritte, so wurde die Textstelle erst mit „Weitergabe an Dritte“ und zusätzlich noch mit „Explizite Nicht-Verarbeitung“ getaggt, wobei hier natürlich der die Negation zum Ausdruck bringende Ausdruck – zum Beispiel „keine“ – ebenso annotiert wurde. Das gleiche Prinzip wurde in ähnlichen Fällen, zum Beispiel „Wenn-Dann-Phrasen“, angewandt.

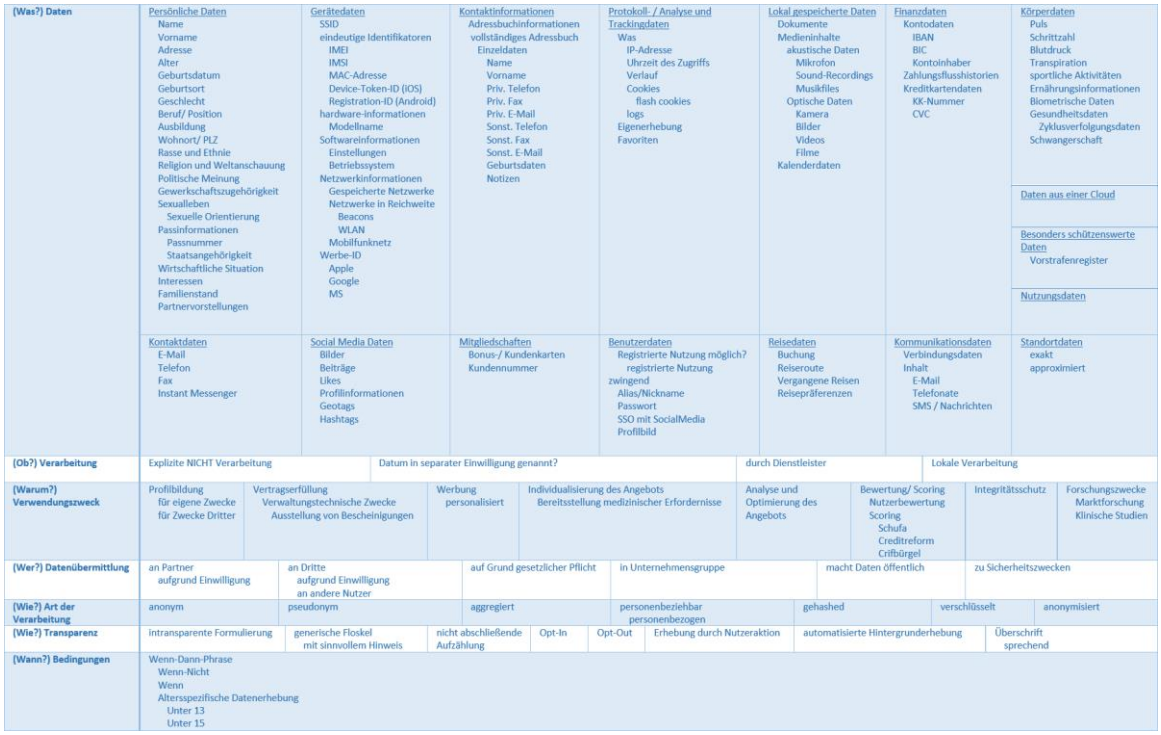


Abbildung 48: Beispielhafte Übersicht gewählter Annotationen hinsichtlich verarbeiteter Daten

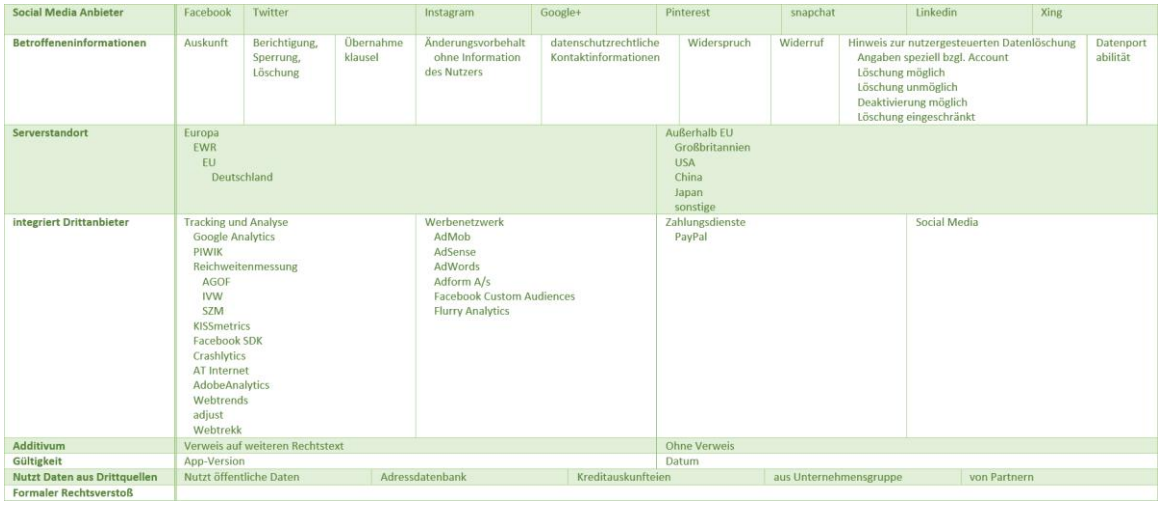


Abbildung 49: Beispielhafte Übersicht über ergänzende Annotationen

Insgesamt wurden 175 Datenschutzerklärungen (dies sind ca. 17.000 Sätze = 270.000 Wörter = 2 Mio. Zeichen) annotiert. Es standen somit ca. 25.000 Annotationen als Trainingsgrundlage zur Verfügung.

4.10.7 Pre-Tagging-Tool

Um das Annotieren von Datenschutzerklärungen zu beschleunigen, wurde ein Active-Learning-System programmiert. Dieses lädt noch nicht bearbeitete Datenschutzerklärungen aus dem Annotationstool herunter und berechnet auf Grundlage der bisher getätigten Annotationen Vorschläge für potentielle Annotationen. Der Annotator muss daraufhin nur noch entscheiden, ob diese potentielle Annotation

richtig oder falsch ist. Anschließend werden positiv bewertete Annotationen an das Annotationstool zurückgeschickt und die Datenschutzerklärung kann für nicht erkannte Annotationen weiter im Annotationstool bearbeitet werden. Je mehr Annotationen in der Vergangenheit gemacht wurden, umso besser werden die Vorschläge über die Zeit und die Zeitersparnis erhöht sich.

Abbildung 50 zeigt eine beispielhafte Ansicht des Pre-Tagging-Tools. Auf der linken Seite befinden sich die verfügbaren Dokumente. Grün markierte Dokumente wurden bereits vollständig bewertet und an das Annotationstool zurück übertragen. Mit einem Klick auf ein Dokument wird dieses in der Mitte angezeigt. Auf der rechten Seite befindet sich nun eine Liste mit Vorschlägen an Annotationen. Fährt der Nutzer mit der Maus über den Vorschlag (hover) so wird die entsprechende Textstelle farblich markiert. Der Annotator entscheidet nun nur noch, ob dieser Vorschlag richtig oder falsch ist

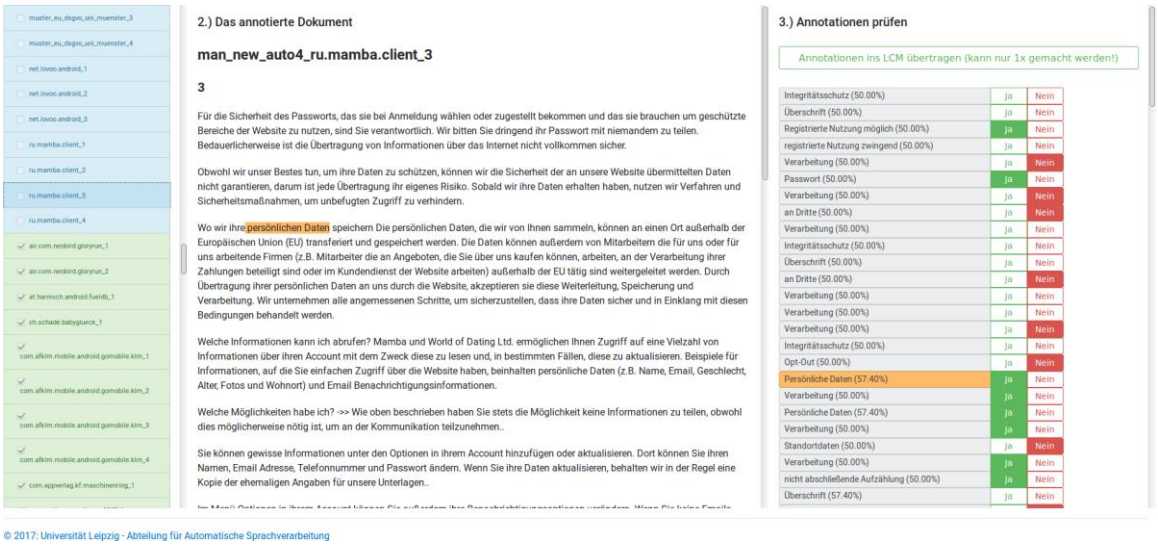


Abbildung 50: Beispielhafte Ansicht des Pre-Tagging-Tools.[57]

4.10.8 Datenschutzrechtliche Kontaktinformationen - Extraktor

Um die Erkennungsrate von datenschutzrechtlichen Kontaktinformationen innerhalb von Datenschutzerklärungen zu erhöhen wurde ein regelbasiertes Programm entwickelt. Dieses erkennt und extrahiert diese Informationen, falls sie in der Datenschutzerklärung vorhanden sind. Dabei kommen hauptsächliche reguläre Ausdrücke und Listen zum Erkennen von deutschen Adressen zum Einsatz. Dies umfasst das Auffinden von:

- Städtenamen (basierend auf Namenslisten vom „Deutschen Wortschatz“[58] und von OpenStreetMap[59])
- Namen von Unternehmen (Schlüsselwörter wie GmbH, Inc., Co. KG.)
- Ländernamen (Liste aller Länder)
- E-Mail-Adressen (regulärer Ausdruck)

[57] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.
[58] http://wortschatz.uni-leipzig.de/de
[59] http://download.geofabrik.de/europe/germany.html

- Personennamen (Namenslisten vom „Deutschen Wortschatz“)
- Straßennamen (reguläre Ausdrücke, Schlüsselwörter, Namensliste von OpenStreetMap)
- Telefonnummern (reguläre Ausdrücke)
- Webseiten (reguläre Ausdrücke)

Für internationale Adressen wird auf „libpostal“[60] zurückgegriffen.

Abbildung 51: Beispielhafte Ausgabe von gefundenen Kontaktdaten innerhalb einer Datenschutzerklärung.

4.10.9 Trainieren von Klassifikatoren

Mit Hilfe der Trainingsdaten (Annotationen) können nun Klassifikatoren für die Analyse von Datenschutzerklärungen erstellt werden. Eine genaue Beschreibung der Trainingsmethode der Klassifikatoren sowie eine Evaluation der Analyseperformance findet sich in Abschnitt 4.10.13. Mit Hilfe der Klassifikatoren können nun unbekannte Texte analysiert werden, um sodann automatisch mit Annotationen versehen werden zu können.

4.10.10 Annotationen zu Infobox Logiken

Die berechneten Annotationen müssen anschließend für den Nutzer zu Informationstexten zusammengefasst werden. Dies geschieht über einfache Logiken, wie zum Beispiel „Falls diese Annotation mindestens einmal gefunden wurde und nicht negiert wurde, zeige folgenden Informationstext an“. Diese Logiken können natürlich auch auf manuell erstellte Annotationen (zum Beispiel aus dem Annotationstool) angewendet werden. Insgesamt gibt es 18 solcher Logiken für die Ergebnisse der semantischen Analyse.

[60] https://github.com/openvenues/libpostal

4.10.11 Backend der semantischen Analyse

Die Ergebnisse der semantischen Analyse einer Datenschutzerklärung zu einer bestimmten App aus dem Google Play Store werden in einer Datenbank (semantisches Backend) gespeichert. Mittels einer API-Schnittstelle werden die Daten für das PGuard-Backend (siehe 4.11.1.2) und das PGuard Browser-Plugin (siehe 6.3) zugänglich gemacht. Die Daten werden dabei in einem definierten JSON Format zurückgegeben (siehe **Fehler! Verweisquelle konnte nicht gefunden werden.**). Sind noch keine Analysedaten zu einer App vorhanden, so wird automatisch nach einer passenden Datenschutzerklärung im Annotationstool gesucht. Soweit keine Daten vorhanden sind, wird zudem im Google Play Store nach einer Datenschutzerklärung gesucht und diese anschließend analysiert. Dafür werden die zuvor beschriebenen Komponenten[61] genutzt.

4.10.12 Privacy Policy Picker

Das „Privacy Policy Picker“ Tool überprüft in regelmäßigen Abständen die Snapshots der Datenschutzerklärung-Webseiten von ca. 30 Services bei archive.org. Die hierbei genutzte und entwickelte Technik kann auch für die Informationsaufbereitung im Rahmen dieses Forschungsvorhabens genutzt werden. Wie in 4.10.5 erläutert erscheint ein Hinweis an die Nutzerinnen und Nutzer sachdienlich, ob abweichende Datenschutzerklärungen im Google Play Store und in der App vorliegen. Ebenso erscheint es sachdienlich Nutzerinnen und Nutzern über eine Lösung wie der im Forschungsprojekt entwickelte Demonstrator des App-Client proaktiv auf Änderungen in den Datenschutzerklärungen der genutzten Apps hinzuweisen.

Über den Privacy Policy Picker können Datenschutzerklärungen über die Zeit hinweg beobachtet und analysiert und etwaige Änderungen angezeigt werden. In einem Webfrontend können ein Service und anschließend zwei Zeitpunkte der Datenschutzerklärung ausgewählt werden. Anschließend wird die Differenz der beiden Versionen dargestellt. Änderungen werden somit schnell und verständlich sichtbar.

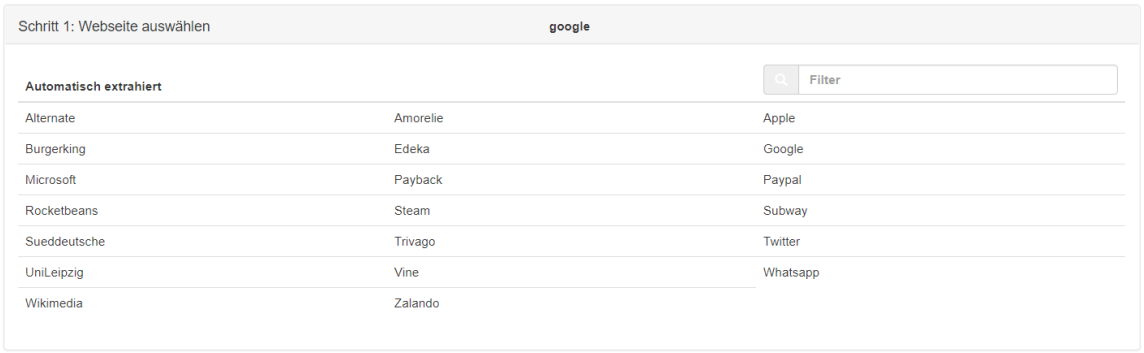


Abbildung 52: Zuerst wählt der Nutzer einen Service bzw. eine Webseite aus, für welche die Datenschutzerklärungen über die Zeit verglichen werden sollen.

[61] Siehe für die Suche 4.10.2, 4.10.3 und 4.10.5; für die Analyse insbesondere 4.10.6 bis 4.10.9.

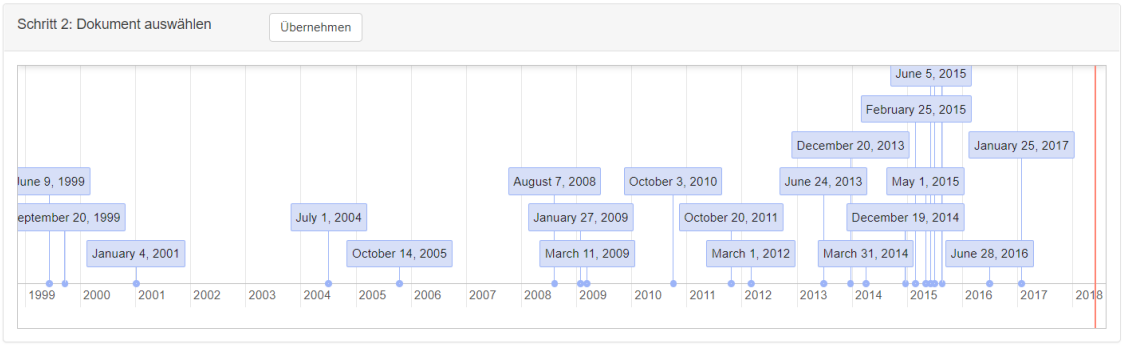


Abbildung 53: Nachdem ein Service ausgewählt wurde, erscheinen alle verfügbaren (textlich verschiedenen) Datenschutzerklärungen auf einem Zeitstrahl. Hier können nun zwei Zeitpunkte ausgewählt werden, um anschließend verglichen werden zu können.

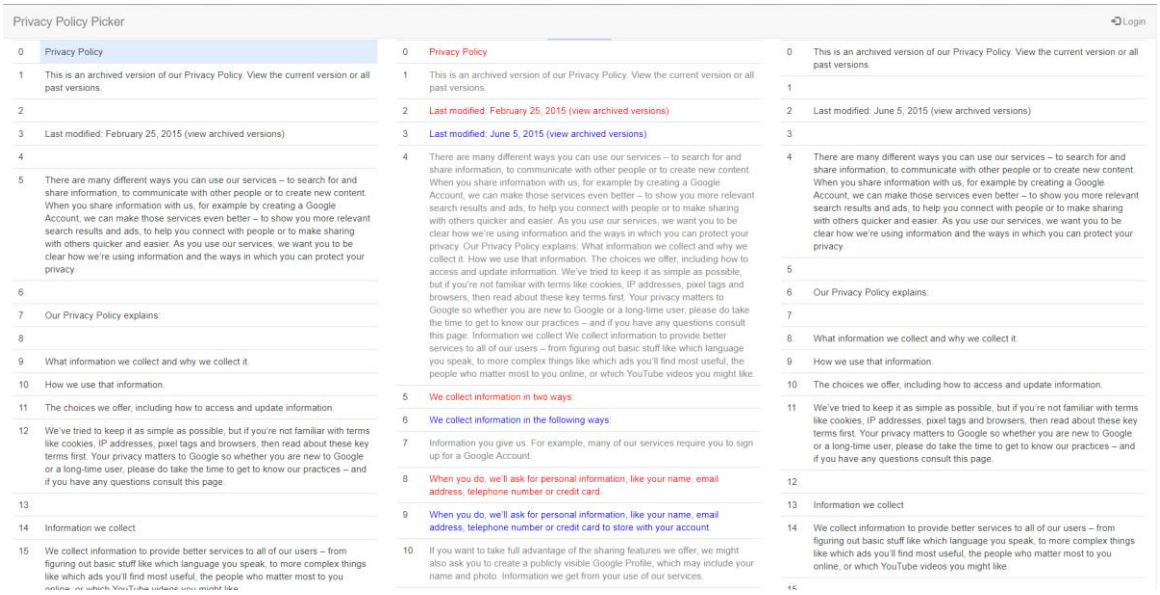


Abbildung 54: Auf der linken und rechten Seite befindet sich jeweils eine Datenschutzerklärung in der Fassung eines bestimmten Zeitpunkts. In der Mitte wird die Differenz angezeigt und wird somit schnell ersichtlich.

4.10.13 Analysemethoden im Detail

Die semantische Analyse von Datenschutzerklärungen wird als Klassifikationsproblem verstanden. Jedem Satz der Datenschutzerklärung wird durch verschiedene Klassifikatoren ein oder mehrere Labels zugeordnet, welche am Ende zum Anzeigen eines Informationstextes führen können (siehe Abbildung 55 und Abbildung 56). Grundlage für das Trainieren der Klassifikatoren sind die manuell erstellten Trainingsdaten (siehe Abschnitt 4.10.6). Für jede Kategorie steht also eine Liste an Trainingssätzen zur Verfügung.

Abbildung 55 stellt schematisch die prinzipielle Aufgabe der semantischen Analyse dar. Die Datenschutzerklärung wird als digitaler Text in die Analyse gegeben. Diese gibt daraufhin als Ausgabe aus, ob bestimmte Informationstexte für diese Datenschutzerklärung anzeigt werden sollen oder nicht.

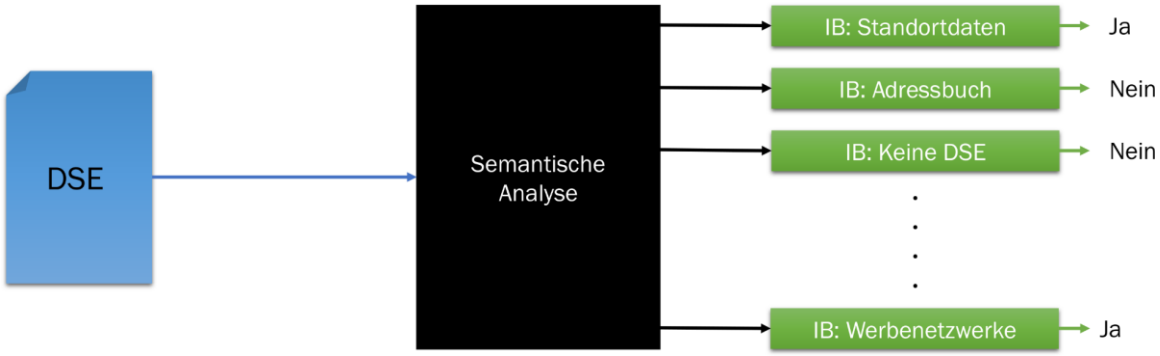


Abbildung 55: Prinzipielle Aufgabe der semantischen Analyse einer Datenschutzerklärung

Abbildung 56 gibt weiteren schematischen Aufschluss, was unter semantischer Analyse zu verstehen ist. Im Forschungsprojekt wurde unter semantischer Analyse die Klassifikation von Sätzen einer Datenschutzerklärung verstanden.

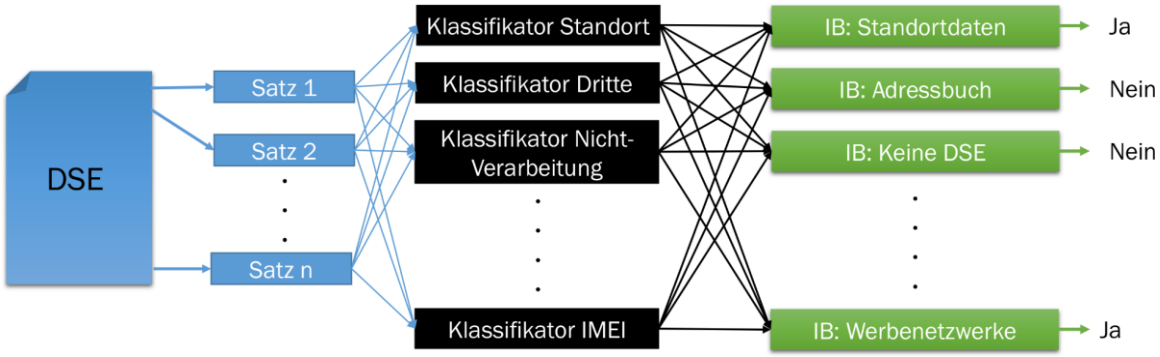


Abbildung 56: Die Blackbox „Semantische Analyse“ aus Abbildung 55 aufgeschlüsselt

Diese Sätze werden, bevor der Klassifikator trainiert wird, vorverarbeitet. Die Vorverarbeitung umfasst folgende Schritte:

1. Es werden alle Ziffern aus dem Text entfernt.
2. Es werden alle Satzzeichen und einige Sonderzeichen entfernt (unter anderem: , . ; : - ? ! _).
3. Der Text wird in Kleinbuchstaben umgewandelt.
4. Es werden, je nach Experiment-Einstellung (siehe unten), die Top-X häufigsten Stoppwörter der deutschen Sprache entfernt.
5. Es werden einzelnstehende Zeichen (also Wörter der Länge 1) entfernt.
6. Es werden einzelnstehende Doppel-Zeichen (also Wörter der Länge 2) entfernt, mit Ausnahme von existierenden deutschen Wörtern wie zum Beispiel in, an, um, …

Für das Training von Klassifikatoren sind sowohl positive als auch negative Trainingsbeispiele nötig. Daher werden zu den positiven Trainingsbeispielen genauso viele negative Beispiele aus den restlichen Daten uniform gesampelt.

Diese positiven und negativen Trainingssätze werden in eine Dokument-Term-Matrix umgewandelt. Jeder Trainingssatz wird als ein Dokument betrachtet. Die „Terme“ können je nach Experimenteinstellung entweder Unigramme, Bigramme, Trigramme oder eine Konkatenation von Uni-, Bi- und Trigrammen sein. Die Gewichtung der Matrix ist auch abhängig von der Experimenteinstellung und kann entweder eine reine Termfrequenz sein oder das Tf-idf-Maß.

Bei der Evaluation wurden mehrere Klassifikationsmethoden und deren Parameter gegeneinander verglichen:

1. SVM
    a. C-Parameter: 0.1, 1, 10
    b. Kernel: Linear, sigmoid
2. Random Forest
    a. Number_of_trees: 63, 255, 511
3. Naive Bayes
    a. LaPlace-Smoothing: 1, 3

Weiterhin wurde verglichen, ob die DTM-Gewichtung (Tf, Tf-idf) einen Unterschied macht. Auch wurde verglichen, ob sich die Anzahl der entfernten Stoppwörter (0, 25, 250) auf die Performance auswirkt.

Als Evaluationsmaße dienen Precison, Recall, F1-Score und Accuracy. Die Berechnung dieser Maße erfolgt mittels einer k-fold Crossvalidation mit k=10.

Das oben beschriebe Experiment wurde 30-mal wiederholt. Dabei änderten sich pro Durchgang nur die uniform gesampelten Negativbeispiele.

Informationstext	num. pos. Trainingssätze	Algorithmus	N-Gram	Terme	DTM-Weighting	Stopword	Accuracy	F1-Score	Precision	Recall
IB 06 – ungenaue Formulierungen	834	SVM-1-linear	uni+bi+tri_grams	58699	weightTf	0	**80.98% (0.71%)**	80.84% (0.80%)	84.73% (0.78%)	77.46% (1.09%)
IB 08 – Standortdaten	311	RF-511	uni_grams	2672	weightTf	250	**87.70% (0.78%)**	86.28% (0.90%)	95.01% (1.26%)	79.36% (1.23%)
IB 09 – DSE Änderung ohne Information	52	NB-1	uni_grams	661	weightTf	250	**93.12% (2.78%)**	92.27% (3.28%)	95.21% (3.13%)	90.89% (4.99%)
IB 10 – Profilergänzung durch öffentliche Daten	13	RF-511	uni_grams	202	weightTfIdf	250	**79.00% (5.48%)**	60.17% (8.29%)	64.33% (9.63%)	59.50% (8.02%)
IB 11 – Verarbeitung durch Dienstleister	215	RF-511	uni_grams	2139	weightTf	25	**87.53% (1.12%)**	86.93% (1.15%)	91.01% (1.66%)	83.66% (1.30%)
IB 12 – App integriert Werbenetzwerke	88	RF-511	uni_grams	1019	weightTf	250	**79.37% (1.50%)**	74.13% (1.82%)	92.43% (2.79%)	63.50% (1.95%)
IB 13 – Erhebung Geräteinformationen	339	RF-255	uni_grams	2782	weightTf	250	**85.19% (1.18%)**	84.51% (1.24%)	88.15% (1.64%)	81.48% (1.44%)
IB 15 – Adressbuch	83	SVM-1-linear	uni_grams	996	weightTf	250	**84.76% (1.69%)**	82.46% (2.10%)	91.55% (2.37%)	76.76% (3.04%)
IB 17 – Übermittlung an Dritte	397	RF-511	uni_grams	3070	weightTfIdf	25	**87.46% (0.75%)**	87.01% (0.73%)	89.51% (1.26%)	84.94% (0.85%)
IB 18 – Erhebung Gerätekennungen	110	SVM-1-linear	uni_grams	1401	weightTf	250	**86.36% (1.74%)**	84.90% (2.10%)	93.31% (1.94%)	79.00% (2.72%)
IB 21 – Zugriff von Drittanbietern	1004	RF-511	uni_grams	5213	weightTf	250	**86.21% (0.50%)**	85.65% (0.49%)	89.12% (0.96%)	82.58% (0.73%)
IB 23 – Teilen in Unternehmensgruppe	71	RF-511	uni_grams	1029	weightTf	250	**86.98% (2.08%)**	85.05% (2.20%)	91.36% (3.55%)	81.31% (2.49%)
IB 27 – Personalisierte Werbung	196	RF-511	uni_grams	2058	weightTf	250	**88.59% (1.31%)**	87.67% (1.37%)	92.38% (1.92%)	83.99% (1.47%)
IB 28 – Verarbeitung im Ausland	82	RF-511	uni_grams	875	weightTf	250	**95.86% (1.04%)**	95.48% (1.08%)	97.80% (1.31%)	93.80% (1.07%)
IB 30 – Macht Daten öffentlich	119	NB-1	uni_grams	1344	weightTf	250	**86.00% (1.17%)**	84.32% (1.47%)	93.32% (2.21%)	77.85% (2.43%)
Metaangaben: explizite Nicht-Verarbeitung (IB 20 – Verarbeitung ausgeschlossener Daten)	469	SVM-1-linear	uni_grams	3064	weightTf	250	**91.55% (0.91%)**	91.68% (0.89%)	90.40% (1.14%)	93.13% (0.93%)

Tabelle 6: Performanceergebnisse der Klassifikatoren für alle relevanten Informationstexte.

In der Tabelle 6 ist für jeden Informationstext der jeweils beste Klassifikator (laut Accuracy) und dessen Parameter aufgelistet. Die Spalte „Informationstext“ gibt den aktuell betrachteten Informationstext an. Die Spalte „num .pos. Trainingssätze“ gibt die Anzahl der positiven Trainingssätze für diesen Klassifikator wieder. Die Spalte „Algorithmus“ gibt den Namen des besten Klassifikationsverfahrens und dessen Parameter an (SVM = Support Vector Machine, RF = Random Forest, NB = Naive Bayes). „N-Gram“ beschreibt die benutzten Features in der Dokument-Term-Matrix. „Terme“ gibt die Anzahl der Terme in der DTM wieder und „DTM-Weighting“ wie die Features in der DTM gewichtet wurden. Die Spalte „Stopword“ beschreibt die Anzahl der entfernten Stoppwörter. Die vier weiteren Spalten geben die Kennzahlen der Evaluation wieder. Die erste Zahl ist dabei jeweils der Mittelwert der Kennzahl über alle 30 Wiederholungen und die Zahl in Klammern die Standardabweichung.

Das beste Ergebnis wurde für „Verarbeitung im Ausland“ erreicht mit ca. 96% Accuracy. Das schlechteste Ergebnis wurde für „Profilergänzung durch öffentliche Daten“ mit ca. 80% Accuracy und nur 60% F1-Score erzielt. Dies liegt daran, dass für dieses Training nur 13 Trainingssätze zur Verfügung standen, was sehr wenig ist. Der Mittelwert der Accuracy über alle Informationstexte liegt bei ca. 87% und der Mittelwert des F1-Scores liegt bei ca. 85%. Somit stellen die Klassifikatoren ein akzeptables Mittel zur semantischen Analyse von Datenschutzerklärungen dar.

Die Performances von SVM’s und Random Forests lagen dabei oft nah beieinander. Aufgrund der Trainingsgeschwindigkeit empfehlen die Autoren SVM als Klassifikationsmethode. Allerdings muss auch festgehalten werden, dass die Trainingsgrundlage mit im Durchschnitt 272 Trainingsbeispielen für quantitative Methoden eher gering ist.

Bei der Klassifikation von unbekannten Texten wird für jede zurückgegebene Annotation die Konfidenz für diese mit angegeben. Die Konfidenz entspricht dabei der Accuracy und dem F1-Score des Klassifikators und dient somit als Qualität der Risikobewertung.

4.10.14 Ausblick / weitere Forschung

Die Ergebnisse der bisherigen Forschung können in nachfolgenden Arbeiten auf drei Arten verbessert werden:

1. Bereitstellen von mehr Trainingsdaten für die bisherigen Verfahren,
2. Anpassen der bestehenden Verfahren um weitere Features und Vorverarbeitungsschritte,
3. Evaluieren von grundsätzlich neuen Methoden.

Für das Bereitstellen von mehr Trainingsdaten kommen folgende Methoden in Frage:

1. Es werden weitere Datenschutzerklärungen von juristischen Experten annotiert.
2. Einholen von Feedback der Nutzer der DATENSCHUTZscanner App, des PGuard Browser-Plugins und der Check your APPS Webseite. Hier werden den Nutzern die entsprechenden Textstellen angezeigt, die zum Anzeigen eines Informationstextes geführt haben. Durch die Implementation eines „das stimmt nicht“-Buttons könnten die Nutzer wertvolles Feedback geben.
3. Trainingsdaten können auch durch Crowdsourcing von Nicht-Experten für einfache juristische Bestandslagen erstellt werden (zum Beispiel durch Beantwortung von einfachen Fragen wie: „Beschreibt dieser Textabschnitt, dass Standortdaten erhoben werden?“).
    a. Geeignete Plattformen könnten sein
        i. UpWork oder
        ii. Amazons Mechanical Turk.
    b. Eine weitere Möglichkeit wäre die Implementation eines „Academic Captchas“-System, indem der Nutzer, anstatt wie bei anderen Captchas Text einzugeben, eine einfache datenschutzrechtliche Frage beantworten muss (zum Beispiel „Welcher dieser 4 Sätze beschreibt den Umgang mit Standortdaten?“)

Weiterhin können die Vorverarbeitungsschritte weiter optimiert werden. Dies würde die spätere Performance positiv beeinflussen. Abgesehen von 1. zielen alle anderen Vorverarbeitungsschritte auf eine Normalisierung des Textes und somit auf eine Verkleinerung des Suchraums ab.

1. Der Datenschutzerklärungstext-Extraktor basiert bisher auf manuell erstellten Regeln. Mittels NLP-Verfahren, wie zum Beispiel Dokumentenklassifikation, könnte die Extraktionsperformance, durch die bessere Generalisierung dieser Verfahren, verbessert werden
2. Abkürzungen in ausgeschriebener Form ersetzen.
3. Stemming der Wörter zu Grundformen
4. Einbinden von bestehendem Weltwissen (zum Beispiel Wordnet, DPPedia, Ontolex, Verbnet) um Wörter durch Synonyme (zum Beispiel „Standortdaten“, „Geodaten“, „Aufenthaltsort“ wird zu Standort) oder Oberbegriffe zu vereinheitlichen.

Die bisher genutzten Klassifikationsverfahren könnten noch wie folgt weiter verbessert werden:

1. Kontext bei der Klassifikation nutzen: nicht nur den aktuellen Satz betrachten, sondern auch den Satz davor und danach.
2. Dependenzgrammatiken benutzen, um einen Satz in Phrasen einzuteilen und diese getrennt zu klassifizieren und Zusammenhänge zwischen den Satzteilen zu erörtern. Dies könnte insbesondere die Performance bei Negation und Wenn-Dann-Bedingungen verbessern.
3. Andere bzw. weitere Features in der Dokument-Term-Matrik nutzen:
    a. Wörter vorher mit POS-Tags anreichern
    b. Word-Embeddings als Features nutzen (gute Generalisierung/Normalisierung von Wörtern)

Die folgende Liste führt weitere Verfahren auf, die gegen die aktuelle Methodik evaluiert werden kann:

- Deep Learning
- Named-Entity-Recognition mit Entity-Linking
- Anstelle von vielen binären Klassifikatoren einen Multiclass-Klassifikator verwenden

4.10.15 Vergleich mit verwandten Arbeiten

Während der Projektlaufzeit wurden folgende ähnliche Projekte zur (semi-)automatischen Analyse im Bereich von Datenschutzerklärungen bekannt.

4.10.15.1 CLAUDETTE meets GDPR

CLAUDETTE[62] analysiert englischsprachige Datenschutzerklärungen auf drei Aspekte: Verständlichkeit, Klarheit und Einhaltung der DSGVO. Das System klassifiziert komplette Sätze innerhalb einer Datenschutzerklärung mittels SVM’s. Andere Verfahren wie Deep Learning (Convolutional Neural Networks) und Long Short-Term Memory Networks wurden ebenfalls evaluiert, erzielten aber schlechtere Ergeb-

[62] http://www.beuc.eu/publications/beuc-x-2018-066_claudette_meets_gdpr_report.pdf

nisse als die SVM's. Für das Training der SVM standen 3.500 annotierte Sätze aus 14 Datenschutzerklärungen zur Verfügung. Die Performance wurde durch ein leave-one-out Verfahren evaluiert und erzielte Werte von 81% Recall und 32% Precision (entspricht 46% F1-Score).

4.10.15.2 Polisis

Polisis[63] benutzt Deep Learning Technologien um Datenschutzerklärungen zu analysieren. Diese benutzen aus 130.000 Datenschutzerklärungen generierte Word Embeddings als Grundlage. Anschließend werden die Datenschutzerklärungen in einem hierarchischen Verfahren klassifiziert. Am Ende werden die Ergebnisse in einem Flowchart (welche Daten werden zu welchen Zwecken erhoben und welche Einstellungsmöglichkeiten gibt es) dargestellt. Es wird eine Accuracy von ca. 88% erreicht. Die Ergebnisse werden in folgenden Papern beschrieben:

- Polisis: Automated Analysis and Presentation of Privacy Policies Using Deep Learning (USENIX Security, 2018)
- PriBots: Conversational Privacy with Chatbots (Twelfth Symposium on Usable Privacy and Security, 2016)

Die Webpräsenz verspricht eine Analyse in allen Hauptsprachen. Allerdings wurden nur englischsprachige Ergebnisse gefunden.

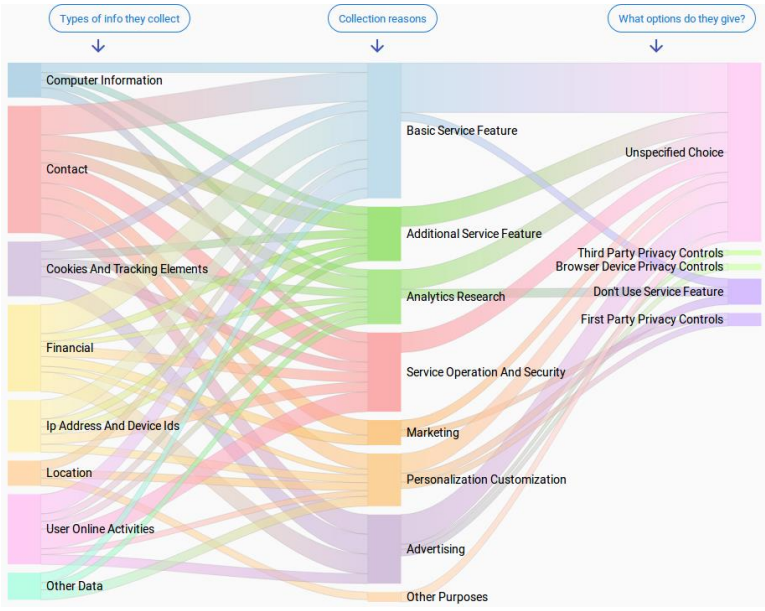


Abbildung 57: Flow-Chart Ausgabe von Polisis.

4.10.15.3 EULAide

EULAide[64] analysiert End-User-Licence-Agreements in englischer Sprache. Die eingesetzten Verfahren basieren auf Ontologien. Die Ergebnisse erreichen F1-Scores im Bereich von 75%.

[63] https://pribot.org
[64] https://github.com/SmartDataAnalytics/EULAide

Die Ergebnisse werden in folgenden Papern beschrieben:

- „EULAide: Interpretation of End-User License Agreements using Ontology-Based Information Extraction“ (SEMANTICS, 2016)
- „Semantic Similarity based Clustering of License Excerpts for Improved End-User Interpretation“ (SEMANTICS, 2017)

4.10.15.4 Open Digital Rights Language – ODRL

Open Digital Rights Language[65] (ODRL) dient einer flexiblen und interoperablen Aufbereitung und Analyse von Nutzungsbedingungen digitaler Inhalte und Dienste. Dieses Modell dient ausdrücklich nicht der Aufbereitung von Datenverarbeitungen hinsichtlich personenbezogener Daten.[66]

Das Modell erinnert stark an P3P[67], allerdings mit einem generischeren Ansatz. Sollte sich eine signifikante Verbreitung von ODRL ergeben und sollte ODRL zukünftig datenschutzrechtliche Aspekte abbilden, könnte dieses Modell im Rahmen der automatisierten Auswertung von Datenschutzerklärungen behilflich sein. Zudem könnten die in ODRL entwickelten Logiken zur Kontextualisierung von Informationen bei einer entsprechenden Weiterentwicklung der semantischen Analyse ein zielführender Anknüpfungspunkt sein.

4.10.16 Zusammenfassung

Für die semantische Analyse galt es zunächst eine technische Infrastruktur zu entwickeln, die die zu analysierenden Texte – nämlich die Datenschutzerklärungen – ermittelt und zugänglich macht. Im Folgenden mussten diese Texte in ein standardisiertes Format übertragen werden.

Die Analyseergebnisse und Zuverlässigkeit waren durchweg gut. Es gibt kein für alle Aspekte hinweg optimales Analyseverfahren, obgleich SVM’s im Rahmen des Forschungsprojekts durchweg positive Ergebnisse erzielten. Selbst je Aspekt kann sich – bei veränderter Datengrundlage – das optimale Analyseverfahren ändern; entsprechend wurde ein System implementiert, mit welchem das optimale Analyseverfahren je Aspekt ermittelt werden kann.

4.11 Zusammenführung der semantischen und technischen Analyseergebnisse sowie Datenfluss bei Informationstexten

Ein Kernaspekt des Forschungsprojektes war es, dass die Prüfergebnisse auf zwei unterschiedlichen Analysen basieren. Hierdurch ist es möglich, die Prüfergebnisse intern zu plausibilisieren sowie – wo möglich und erforderlich – zu kumulieren.

Hierzu galt es zunächst eine technische Infrastruktur zu entwickeln, die beide Analyseergebnisse zusammenführt und eine technische beziehungsweise semantische Analyse einer App im Bedarfsfall auslösen kann.

[65] https://www.w3.org/TR/odrl-model/.
[66] https://www.w3.org/TR/2018/REC-odrl-model-20180215/#privacyConsiderations.
[67] https://www.w3.org/P3P/.

Auf Basis dieser zentralen Erkenntnisse galt es zudem, die relevanten Erkenntnisse für die Informationstexte zu ermitteln und in eine Vereinigungslogik zu überführen.

4.11.1 Zusammenführung der semantischen und technischen Analyseergebnisse

Zur Integration der beiden Informationsquellen und der Harmonisierung der Datenbestände wurden mehrere Komponenten entwickelt, die eine Integration in verschiedene zukünftige Systeme ermöglicht.

4.11.1.1 PGuard Web-Backend

Die im PGuard-Backend[68] zusammenlaufenden Ergebnisse galt es auch anschaulich für die weitere Bearbeitung innerhalb des Projektes zugänglich, prüfbar und zum Teil bearbeitbar zu machen. Hierzu wurde das PGuard-Web-Backend entwickelt. Das System umfasst mehrere Ansichten der analysierten Apps und der jeweils zur Verfügung stehenden Testergebnisse. Zunächst gibt es die harmonisierte Ansicht inklusive der Metadaten (Abbildung 58).

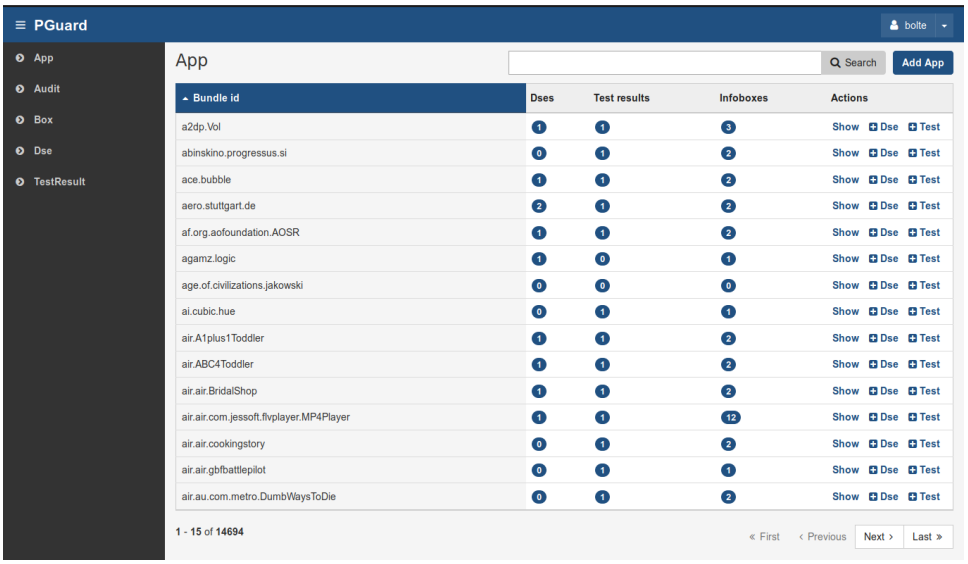


Abbildung 58: Backend– Grafische Oberfläche mit verfügbaren Informationen[69]

Weiterhin gibt es eine Ansicht mit detaillierten Infos aus der Datenschutzerklärungsanalyse und den dazu gehörenden Metadaten (Abbildung 59). Zusätzlich gibt es eine Ansicht der Details aus den Technischen- und Malwareanalysen (Abbildung 60).

[68] Siehe 4.11.1.2.
[69] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.

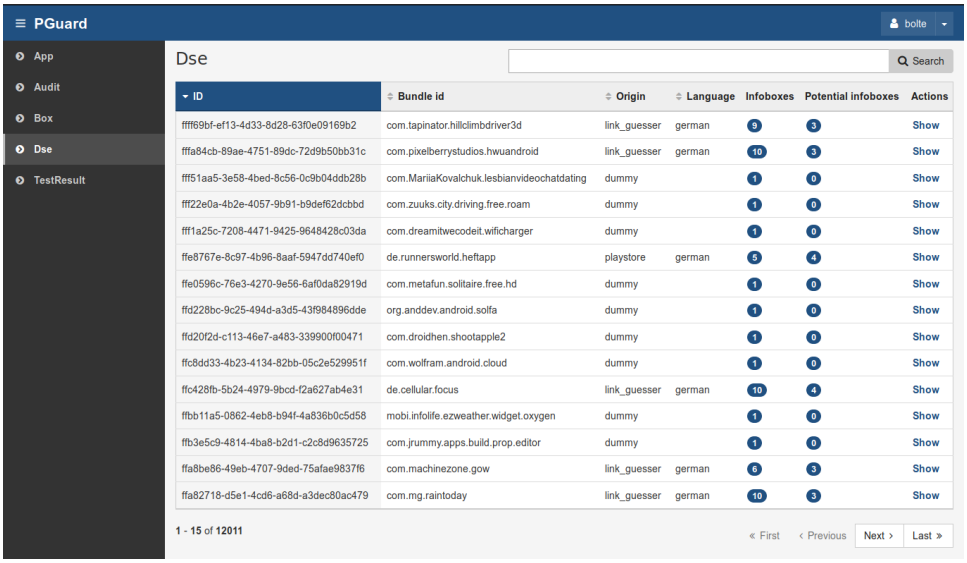


Abbildung 59: Backend– Details zu den gesammelten Informationen aus der Datenschutzerklärungs Analyse[70]

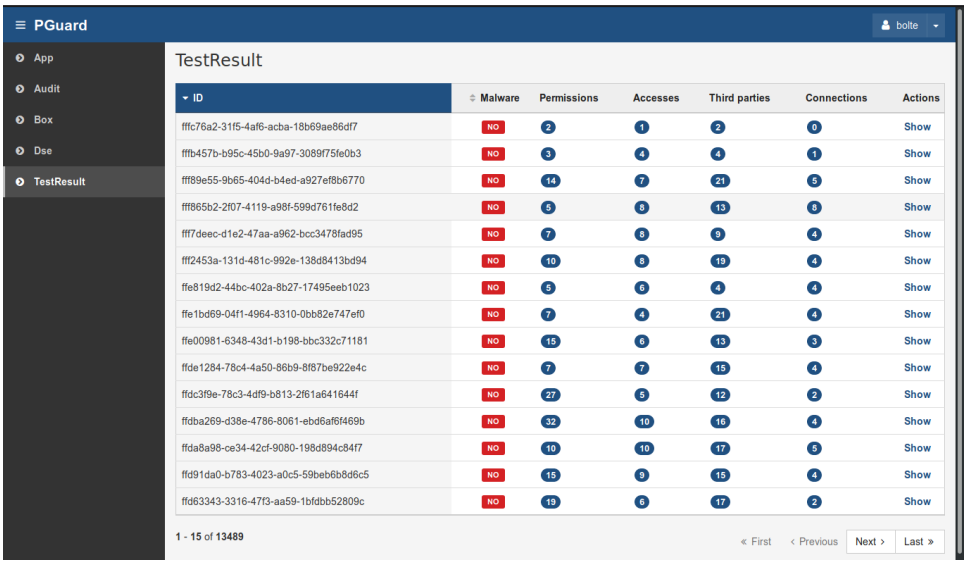


Abbildung 60: Backend– Details zu den gesammelten Informationen aus der technischen Analyse[71]

4.11.1.2 PGuard Backend

Die Zusammenführung von Datensätzen zur Bewertung der technischen und semantischen Prüfergebnisse einer App wurde in einem eigenen Backend implementiert (PGuard-Backend). Hierbei wurde – soweit sachdienlich – auf bestehende Lösungen zurückgegriffen und das PGuard-Backend flexibel ausgestaltet. So werden die jeweiligen Prüfergebnisse per API Schnittstellen an das PGuard-Backend übertragen. Die Verarbeitungslogiken zur Bewertung der Ergebnisse findet sodann im PGuard-Backend statt.

[70] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.
[71] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.

Hierdurch ist es möglich, für die Prüfergebnisse auf Dauer die Quellen auszutauschen als auch um weitere Quellen anzureichern. Das dedizierte PGuard-Backend kann somit eigene Anfragen zur Analyse von mobilen Applikationen in den jeweiligen semantischen und technischen Backends initiieren. Die auf Basis dieser Anfragen generierten Analyseergebnisse werden sodann ebenfalls via API an das PGuard-Backend übertragen und in einer eigenen Datenbank abgelegt.

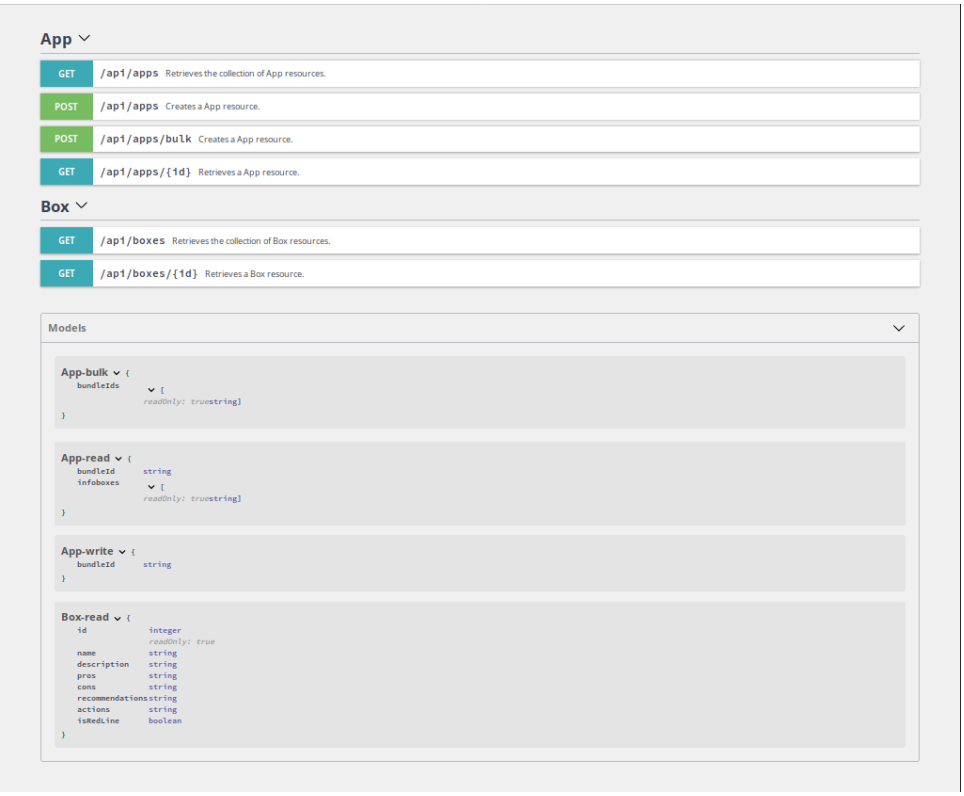


Abbildung 61: API– Selbstpflegende API Dokumentation

Die im PGuard-Backend zusammengefassten und analysierten Prüfergebnisse können via API an entsprechende Interfaces (wie zum Bespiel den App-Client) ausgeliefert werden. Hierdurch konnten die Prüfergebnisse und die damit verbundenen Informationen in Labormustern, beispielsweise dem DATENSCHUTZscanner[72], dargestellt werden.

Damit die Details zu den Ergebnissen ebenfalls dynamisch erweitert werden können wurde die API zur Kommunikation mit den Clients dynamisch ausgestaltet, sodass in Zukunft zusätzliche Details hinzugefügt werden können. Hierzu nachstehendes Beispiel:

[72] Siehe 5.

```
[
  {
    "id": "8",
    "is_red_line": "false",
    "titel": "Die App nutzt Standortdaten",
    "description": "Die App erfasst Ihre Position. Dies kann sowohl durch GPS als auch die Auswertung des Netzwerkempfangs Ihres Mobilfunkbetreibers oder WLAN-Signals geschehen.",
    "pros": [
      {
        "first_layer": "Standortbezogene Funktionen können genutzt werden, ohne dass Sie den Standort extra eingeben müssen.",
        "second_layer": "Beispiele sind Navigation und Umgebungssuche oder standortbasierte Werbung."
      }
    ],
    "cons": [
      {
        "first_layer": "Ein dauerhafter Zugriff ermöglicht die Erstellung von Bewegungsprofilen.",
        "second_layer": ""
      }
    ],
    "recommendations": "Aktivieren Sie nur die standortbezogenen Funktionen Ihres Smartphones, die zum jeweiligen Zeitpunkt benötigt werden (GPS, WLAN, Bluetooth, etc.). Wenn Sie das nicht möchten, verbieten Sie der App den Zugriff auf den Standort. Gewähren Sie der App nur Zugriff auf den genauen Standort (GPS), wenn dies wirklich erforderlich ist. Wenn Sie das nicht möchten, deaktivieren Sie die Standort-Funktion.",
    "actions": [
      "Zugriff verbieten auf den individuellen Standort",
      "Deaktivierung der globalen Standortfunktion",
      "Deinstallieren der App"
    ]
  },
]
```


Abbildung 62: JSON Format – Beispielübertragung eines IBs (08 - nutzt Standortdaten)

Hier wird ersichtlich, wie die Daten strukturiert wurden und dass zu jedem IB alle relevanten Informationen über die API ausgeliefert werden können. Diese Informationen umfassen:

- id: Informationsblock Nummer
- is_red_line: handelt es sich um einen Rechtsverstoß?
- titel: Der Titel des IBs
- description: eine Beschreibung der Fundgruppe
- pros: Die Vorteile für den Betroffenen
- cons: Die Nachteile für den Betroffenen
- first_layer: Erweiterte Informationen die zusätzlich zur Verfügung gestellt werden können
- second_layer: Tiefergehende Erweiterte Informationen die zusätzlich zur Verfügung gestellt werden können
- recommendations: Handlungsempfehlungen, damit der Betroffene besser versteht welchen Einfluss der Fund auf seine Privatsphäre hat
- actions: Handlungsempfehlungen für den Betroffenen die auf seinem Endgerät durchgeführt werden können

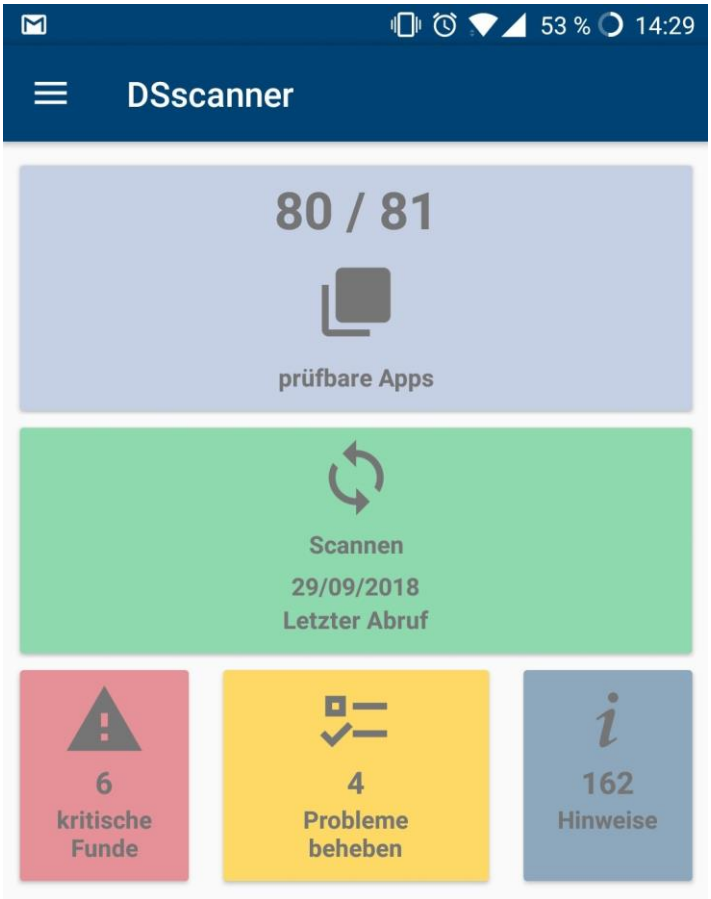


Abbildung 63: App Frontend – Dashboard der App

Sollten zu einer angefragten App keine Ergebnisse vorliegen, werden die notwendigen Analysen nun angestoßen und stehen bei der nächsten Anfrage eines Clients im PGuard-Backend zur Verfügung.

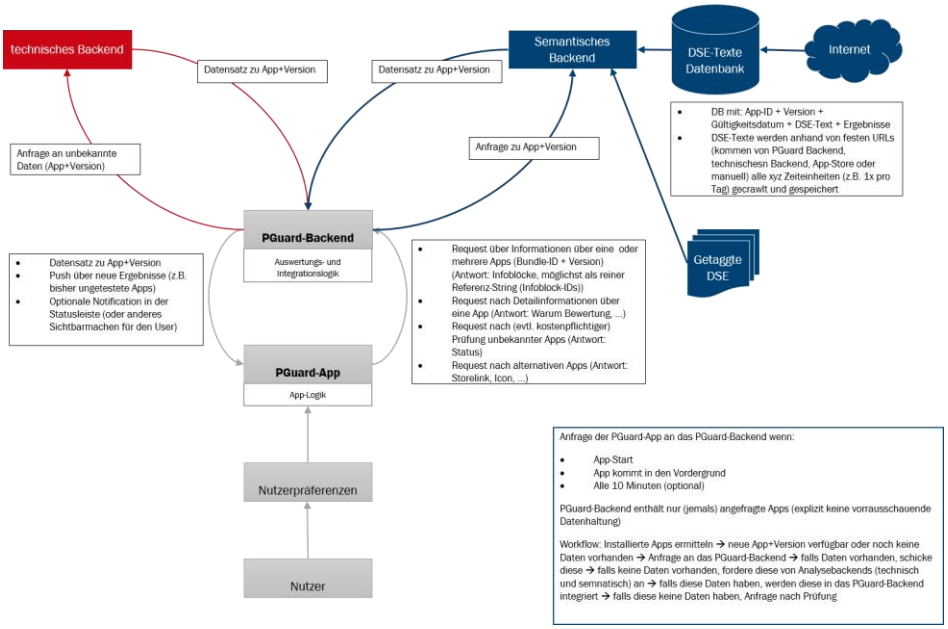


Abbildung 64: Technisches Schema – Backend Integration

4.11.2 Datenfluss bei Informationstexten

Für die Informationsintegration zwischen Daten aus der semantischen Analyse und Daten aus der technischen Analyse wurde auf die unter 4.8.2 beschriebenen Informationstexte zurückgegriffen, welche um die Datenquelle und die Priorisierung der Testergebnisse erweitert wurde, wie im folgenden Beispiel zu sehen ist.

Aus der Darstellung (Abbildung 65) lässt sich folgende Verarbeitung der Daten herleiten:

Die im Informationstext (IT) 08 behandelte Nutzung von Standortdaten hat die aufgezeigten möglichen Vor- und Nachteile. Die Ergebnisse, ob dieser IT zutrifft, ergeben sich aus den technischen und den semantischen Analysen. Daher werden die Informationen aus diesem Informationstext dargestellt, wenn eine der beiden Prüftechnologien die Verarbeitung der Daten feststellt.

<table>
<tr><th colspan="2">IT 08 - Die App nutzt Standortdaten</th></tr>
<tr><td>⊕ Standortbezogene Services können personalisiert genutzt werden, ohne dass Sie den Standort extra eingeben müssen.
⊕ Beispiele sind Navigation und Umgebungssuche oder standortbasierte Werbung.</td><td>⊖ Ein dauerhafter Zugriff ermöglicht die Erstellung von Bewegungsprofilen.</td></tr>
<tr><td>Technische Daten</td><td>Semantische Daten</td></tr>
<tr><td>Aus der automatischen Analyse</td><td>Tags:
• Standortdaten</td></tr>
<tr><td colspan="2">Prüflogik
Antwort: Ja / Nein</td></tr>
<tr><td colspan="2">Vereinigungslogik
<table>
<tr><td colspan="2"></td><td>Technisch
Ja</td><td>Nein</td><td>k.A.</td></tr>
<tr><td>Semantisch</td><td>Ja</td><td>Ja</td><td>Ja</td><td>Ja</td></tr>
<tr><td></td><td>Nein</td><td>Ja</td><td>Nein</td><td>Nein</td></tr>
<tr><td></td><td>k.A.</td><td>Ja</td><td>Nein</td><td>k.A.</td></tr>
</table></td></tr>
</table>

Abbildung 65: Prüf- und Vereinigungslogik am Beispiel eines Informationstextes

4.11.3 Zusammenfassung

Im Rahmen des Projektes wurde ein dynamisches System entwickelt, wie die unterschiedlichen Informationsquellen und Aufbereitungen an zentraler Stelle verwaltet werden können. Hierbei wurde großer Wert auf Flexibilität und Interoperabilität gelegt. Zudem wurden für jeden Informationstext Entscheidungslogiken entwickelt, auf deren Basis ein Informationstext angezeigt wird.

5 Forschung und Ergebnisse im Zusammenhang des App-Clients

Dieses Kapitel widmet sich den Forschungsergebnissen und Herausforderungen im Rahmen der Entwicklung eines App-Clients. Hierbei galt es sowohl generelle Aspekte zu betrachten als auch sehr spezifische Aspekte, die auch mit rechtlichen Implikationen einhergingen.

5.1 Technische Besonderheiten insbesondere hinsichtlich der Umsetzbarkeiten von Handlungsoptionen

Die im Forschungsprojekt angestrebten Funktionen eines App-Clients galt es auch technisch zu evaluieren und umzusetzen.

5.1.1 Erkenntnisgewinn auf technischer Basis

Technische Unterschiede und somit veränderte Entwicklungsbedarfe ergeben sich nicht nur im Rahmen der Generierung der Prüfergebnisse, sondern auch im Rahmen der Entwicklung eines App-Clients.

Dabei sind die Gründe unterschiedlich: Funktionen, Bedienelemente beziehungsweise Bedienoptionen werden durch ein Betriebssystem angeboten, durch ein anderes aber unter Umständen nicht. Auch die Designvorgaben können sich unterscheiden, was letztlich Auswirkungen auf die erwarteten Klickstrecken der Betroffenen haben kann.

Als Beispiel für sich erheblich unterscheidende Ansätze zwischen den Betriebssystemen kann das Auslesen der auf einem Endgerät installierten Applikationen genannt werden: Eine Android-App kann grundsätzlich weitere auf dem Endgerät installierte Applikationen erkennen und zum Teil auch mit diesen interagieren. Eine iOS-App verfügt grundsätzlich nicht über die Möglichkeit sonstige auf einem Endgerät installierte Applikationen zu erkennen, mithin kann eine solche Applikation auch grundsätzlich nicht mit diesen interagieren; näheres hierzu unter 5.3.2.

5.1.2 Handlungsoptionen für Betroffene

Die Handlungsoptionen für Betroffene stellen eine Kernfunktion des App-Clients dar. Die Handlungsoptionen können sich indessen je nach Betriebssystem unterscheiden. Im Forschungsprojekt wurde sich für die Entwicklung eines Android-Clients entschieden. Hierbei wurde als Mindestversion Android 6.0 vorausgesetzt. Hiermit konnte sichergestellt werden, dass alle Handlungsoptionen, die eine Einschränkungen von Zugriffsrechten beinhalten von den Betroffenen auch tatsächlich umgesetzt werden können; näheres hierzu unter 4.6.

5.1.3 Zugänglichkeit und technische Bedienbarkeit der Handlungsoptionen für Betroffene

Die nutzerfreundliche Ausgestaltung der Handlungsempfehlungen war Kern der Überlegungen und Forschung. Hierbei sollten Betroffene möglichst ohne unnötige Zwischenschritte in die Lage versetzt werden, die gewünschte Handlungsoption umzusetzen.

Bei der technischen Ausgestaltung galt es somit bestehende Lösungen einzusetzen beziehungsweise neue Lösungsansätze zu entwickeln, um die Zwischenschritte der Betroffenen zu minimieren.

So werden Betroffene, sollten diese Zugriffsrechte für eine App einschränken wollen, durch den App-Client direkt in das entsprechende Menü geführt. Nutzer müssen in den Einstellungen der App sodann nur noch den Bereich „Berechtigungen“ öffnen und das betroffene Zugriffsrecht deaktivieren. Ein direkter Verweis auf diese letzte Ebene konnte technisch nicht realisiert werden. Es ist allerdings möglich, Betroffenen stattdessen temporär einen Text einzublenden, der den letzten Schritt nochmals erläutert. Letzteres wurde technisch im Demonstrator nicht umgesetzt, da hierzu bereits Umsetzungen in anderen Apps existieren, sodass diese Technik grundsätzlich als funktional betrachtet werden konnte und die Integrationsaufwände mit dem für das Forschungsvorhaben intendierten Ziele nicht im Verhältnis standen.

Im Forschungsprojekt als sinnvoll eingestuft wurden zudem Überprüfungen, ob Betroffene etwaige Zugriffsrechte einer App bereits eingeschränkt haben. Derartige Informationen konnten ohne Weiteres jedoch nicht ausgelesen werden. Ermittelte technische Ansätze hätten sogenannte „root-Rechte“ der Betroffenen auf dem Endgerät erfordert. Da die Verbraucherinnen und Verbraucher voraussichtlich nicht über die notwendigen Rechte verfügen, wurden diese Ansätze nicht weiterverfolgt.

Entsprechendes gilt für das Auslesen und verlinken in andere Apps und deren gegebenenfalls vorhandene Einstellungsoptionen. Es gibt keine generelle Möglichkeit in andere Apps auf eine konkrete Unterfunktion beziehungsweise Unterseite zu verlinken. Hierzu bedarf es grundsätzlich entsprechender, individuell von einer App bereitgestellter Schnittstellen. Technische Ansätze, die dennoch theoretisch eine ähnliche Funktionalität hätten bieten können, wären nicht mit den Appstore-Richtlinien vereinbar; die technische Grenze zur Malware wäre fließend, sodass diese Ansätze ebenfalls nicht weiterverfolgt wurden.

5.1.4 Zusammenfassung

Die Funktionen des App-Clients sind grundsätzlich betriebssystemunabhängig. Hierbei ist aber zu beachten, dass sich deren Implementierung je Betriebssystem technisch unterscheiden wird. In Einzelfällen können Funktionen nur in stark abgewandelter Form realisiert werden, welches auch Auswirkungen auf die Nutzerfreundlichkeit haben kann.

5.2 Besonderheiten hinsichtlich der Zugänglichkeit, Verständlichkeit und Bedienbarkeit eine App-Clients

5.2.1 Umsetzung und übergeordnete Entwicklungsschritte

Im Folgenden werden die Besonderheiten hinsichtlich der Zugänglichkeit, Verständlichkeit und Bedienbarkeit des App-Clients beschrieben. Hierbei wird auf die visuelle Entwicklung der Nutzeroberfläche in der App eingegangen und werden die Ergebnisse aus Nutzertests vorgestellt.

Bei der Entwicklung der Nutzeroberfläche und des Bedienkonzepts wurden insgesamt siebe Schritte absolviert:

1. Schritt: Mission Vision des App-Clients und Usability-Leitplanken
2. Schritt: Mock Ups: Erstellung des groben Bedienkonzepts
3. Schritt: Prototyp: Erweiterung des Bedienkonzepts und optimierte Visualisierung
4. Schritt: Labormuster 2017: Version eines funktionalen Prototyps
5. Schritt: User-Tests auf der Internationalen Funkausstellung Berlin
6. Schritt: Experten Beta-Phase
7. Schritt: Labormuster 2018: Version des weiterentwickelten Prototyps

5.2.2 Mission und Vision des App-Clients und Usability-Leitplanken

Wie bereits in 4.8.1 beschrieben, war es für die Konzeptionierung des App Clients notwendig, klare Leitplanken zu definieren, die beschreiben, welche Mission bzw. Vision die Anwendung erfüllen soll. Die Definition der Mission und Vision lautet:

Der DATENSCHUTZscanner bietet Nutzerinnen und Nutzern mehr Transparenz und Kontrolle beim Thema Datenschutz in ihren Apps.

Um diese Mission und Vision zu erfüllen, war es somit notwendig, Transparenz und Kontrolle als Funktionen und in der Nutzeroberfläche zu berücksichtigen und die Gestaltung des App-Clients hiernach auszurichten.

Über die übergeordnete Ausrichtung des Clients hinaus war es zusätzlich notwendig, die Bedienbarkeit (Usability) zu gewährleisten. Deshalb wurden bei der gesamten Entwicklung und Optimierung des App-Clients sieben aus der Usability-Literatur identifizierte „Goldstandards“ bzw. Eigenschaften berücksichtigt[73]:

1. Die App bzw. angebotene Funktionen in der App sollten **nützlich** sein, das heißt Nutzerinnen und Nutzer ziehen einen Mehrwert aus ihnen.
2. Die App bzw. angebotene Funktionen in der App sollten **erlernbar** sein, das heißt Nutzerinnen und Nutzer sind in der Lage, eigenständig und einfach herauszufinden wie diese funktioniert.
3. Die App bzw. angebotene Funktionen in der App sollten **einprägsam** sein, das heißt nach einmaligem Verstehen muss die App / Funktion nicht erneut erlernt werden.
4. Die App bzw. angebotene Funktionen in der App sollten **effektiv** sein, das heißt wirksam sein.
5. Die App bzw. angebotene Funktionen in der App sollten **effizient** sein, das heißt in einem verhältnismäßig geringen (zeitlichen) Aufwand zum Ziel führen.
6. Die App bzw. angebotene Funktionen in der App sollten **begehrenswert** sein, das heißt von Nutzerinnen und Nutzern als attraktiv empfunden werden
7. Die App bzw. angebotene Funktionen in der App sollten **reizvoll** sein, das heißt Nutzerinnen und Nutzer haben bei der Bedienung Freude und Spaß.

Immer wenn neue Funktionen oder Bedienkonzepte entwickelt wurden, wurden diese mit den oben genannten Eigenschaften abgeglichen.

[73] Krug, S. (2014). Don’t make me think! Revisited. Web & Mobile Usability; S. 9

5.2.3 Mock Ups: Erstellung des groben Bedienkonzepts

Im ersten Schritt zum Projektbeginn wurden sog. Mock Ups, das heißt Attrappen, erstellt, die die unterschiedlichen App-Konzepte visualisieren. In den Mock Ups, die mit dem Programm „Ninja Mock“ erstellt wurden, waren folgende erste Bedienkonzepte und Funktionen enthalten:

a) Nutzerinnen und Nutzer können ihre Präferenzen in Form eines Profils angeben

Diese Funktion soll den Präferenzabgleich ermöglichen und so das übergeordnete Ziel der Anwendung „Kontrolle“ unterstützen. Die Herausforderung in der Umsetzung des Kontrollziels besteht grundsätzlich darin, dass keine Vorbewertung von Apps vorgenommen werden soll, sondern es Nutzerinnen und Nutzern freibleiben soll, wie sie persönlich und gemäß ihren eigenen Präferenzen eine Datenverarbeitung bewerten. Das in den Mock Ups dargestellte „Profil“ soll die Angabe der Präferenzen abbilden (Abbildung 69).

b) Es gibt unterschiedlich kritische Funde

In den ersten Mock Ups wurde das Konzept der „roten Linien“ (vgl. 4.8.2.2) bereits angedacht, jedoch unter dem Arbeitstitel „kritische Funde“ geführt. Diese kritischen Funde sollen Nutzerinnen und Nutzer Datenverarbeitungen anzeigen, die beispielsweise aufgrund von Gesetzesverstößen, negativ vorbewertet sind und so Handlungen zum Selbstdatenschutz motivieren. Die kritischen Funde stehen nicht im Konflikt mit der unter a) genannten Präferenzangabe, da sie lediglich einen kleinen Teil der Datenverarbeitungen betreffen, das heißt bei der Mehrzahl der Datenverarbeitungen liegt die Bewertungsentscheidung bei den Nutzerinnen und Nutzern selbst und nur bei wenigen roten Linien nimmt die App eine Vorbewertung vor.

c) Übersicht über alle Apps

Eine weitere Funktion, die bereits in den Mock Ups berücksichtigt wurde, ist die App-Übersicht. Sie soll Nutzerinnen und Nutzern ermöglichen, Vergleiche zwischen Anwendungen zu ziehen und Datenverarbeitungen im Kontext zu verstehen. Diese Funktion zahlt somit auf das übergeordnete Transparenzziel der Anwendung ein.

Die Abbildungen Abbildung 66 bis Abbildung 69 zeigen Ausschnitte aus den erstellten Mock Ups.

Abbildung 66: Mock Up (1/4)

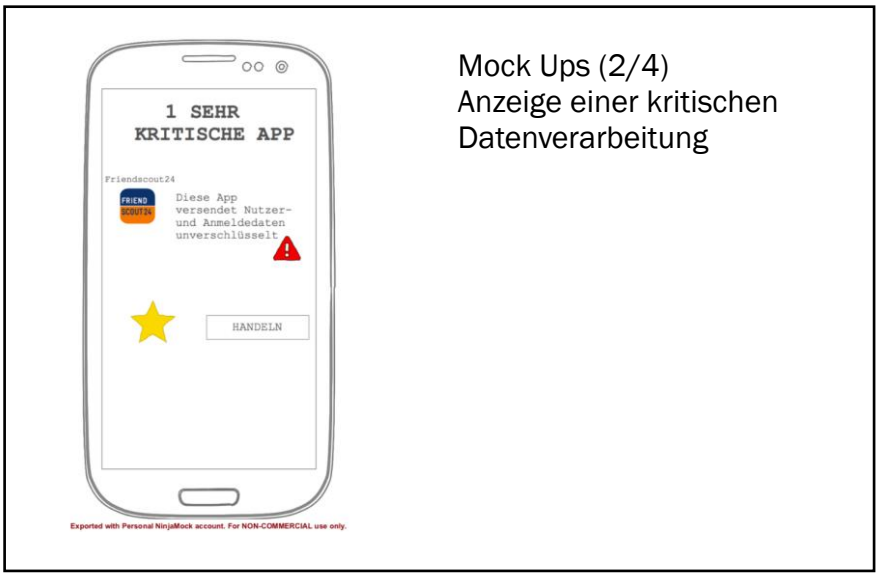


Abbildung 67: Mock Up (2/4)[74]

[74] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.

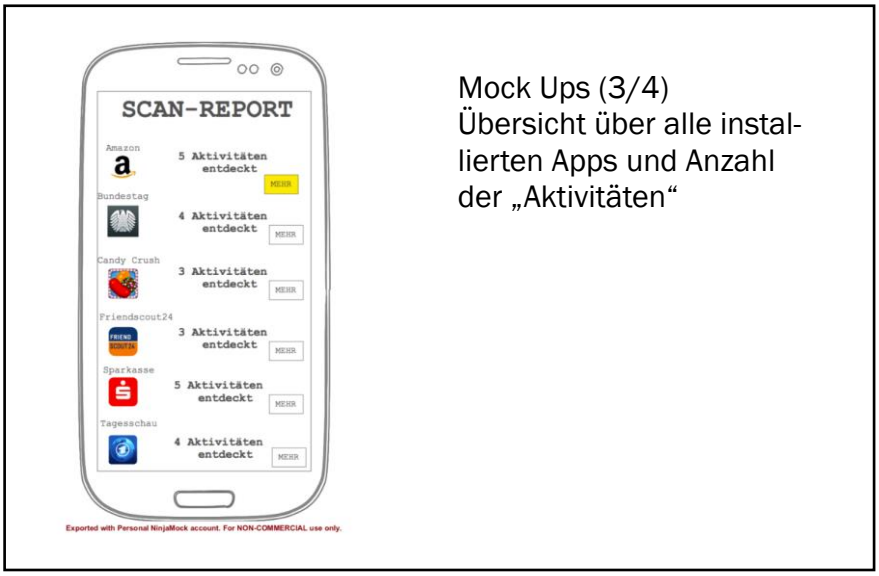


Abbildung 68: Mock Up (3/4)[75]

Abbildung 69: Mock Up (4/4)[76]

5.2.4 Prototyp: Erweiterung des Bedienkonzepts und optimierte Visualisierung

Im nächsten Schritt wurden die Mock Ups durch einen Prototyp erweitert, der die Funktionen und Bedienelemente in optimierter Darstellung abbildet. Hierzu wurde das Programm „Proto.io“ verwendet, das ein Toolkit für typische App-Oberflächen beinhaltet und realen Anwendungen sehr ähnlich ist.

Folgende Funktionen wurden im Prototyp hinzugefügt oder erweitert:

[75] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.
[76] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.

a) Nutzerinnen und Nutzer können kontextspezifisch ihre Präferenzen angeben

Das Bedienkonzept sieht vor, dass Funde (das heißt festgestellte, relevante Datenverarbeitungen) „bearbeitet“ werden. Hierbei können die Nutzerinnen und Nutzer kontext-spezifisch bewerten, ob eine bestimmte Datenverarbeitung mit ihren Präferenzen übereinstimmt. Die Bewertungsfunktion ist somit eine Erweiterung der Profilfunktion aus den Mock Ups und soll zur Erreichung des Kontrollziels implementiert werden.

b) Organisation der Datenverarbeitungen nach Clustern

Eine weitere Funktion, die auf dem Konzept der Informationstexte-Cluster basiert (vgl. 4.8.3.6), ist die Sortierung der Datenverarbeitung. Nutzerinnen und Nutzer erhalten hierbei eine Übersicht darüber, in welchen Clustern Datenverarbeitungen festgestellt wurden und können die Bearbeitung, wie in a) beschrieben, für einzelne Cluster vornehmen, die sie besonders interessieren.

Die Folgenden Abbildungen Abbildung 70 bis Abbildung 72 zeigen einen Ausschnitt aus dem erstellten Prototyp:

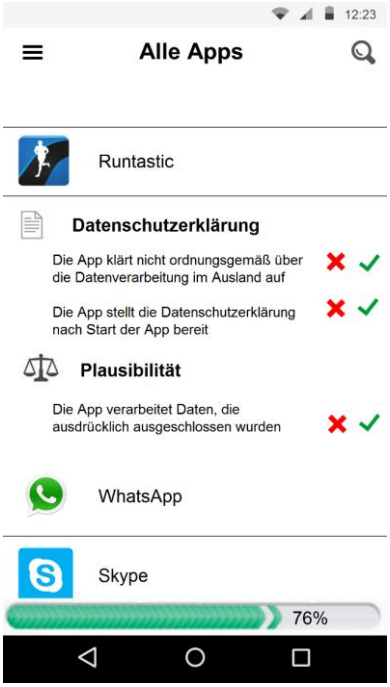


Abbildung 70: Prototyp – Funde zu einer bestimmen App, hier „Runtastic“[77]

[77] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.

Abbildung 71: Prototyp – Übersicht Funde

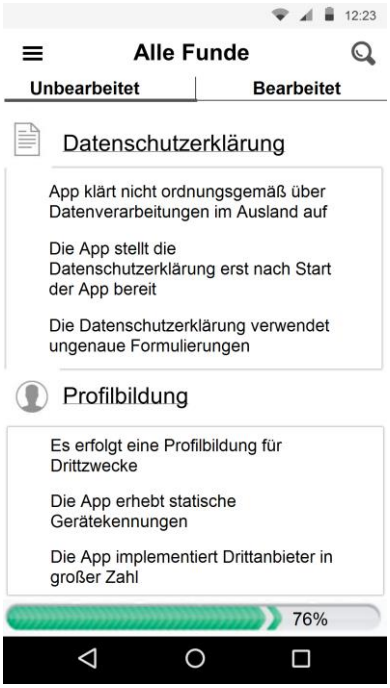


Abbildung 72: Prototyp – Anzeige aller Fundgruppen und deren untergeordneter Funde

5.2.5 Labormuster 2017: Version eines funktionalen Prototyps

Auf Grundlage des Bedienkonzepts, das bereits in den Mock Ups und dem Prototyp visualisiert wurde, wurde das erste Labormuster als Android-App erstellt. Hierfür wurden Funktionen weiterentwickelt und neue Funktionen hinzugefügt.

Im Folgenden werden die Funktionen des Labormusters zusammengefasst:

a) Übersicht über alle Apps, alle Funde und Funde nach Clustern

Basierend auf den Überlegungen zu den Mock Ups wurde an der Funktion, einen Überblick über alle Apps und die Anzahl der festgestellten Datenverarbeitungen (Funde) zu erhalten, festgehalten. Auch die Sortierfunktion nach Funden und Clustern wurde aus dem Prototyp für das Labormuster übernommen.

b) Visualisierung der Informationstexte

Um das übergeordnete Transparenzziel der Anwendung zu erreichen, wurde in das Labormuster ein Design zur Darstellung der Informationstexte integriert. Hierdurch ist es möglich, Nutzerinnen und Nutzern sowohl Informationen zur Datenverarbeitung an sich, als auch zu den Konsequenzen bereitzustellen. Die Umsetzung orientiert sich hierbei an den Vor- und Nachteilen einer spezifischen Datenverarbeitung, so wie sie in 4.8.2.1 bereits dargelegt wurde. Eine weitere Funktion, die in das Labormuster integriert wurde, sind Pop Ups, die in kleinen Informationsfenstern bei Bedarf weiterführende Informationen bzw. Erläuterungen zu Datenverarbeitungen bieten. Diese setzen die sogenannten Detail-Informationen der Informationstexte visuell und funktional um.

c) Nutzerinnen und Nutzer können kontextspezifisch ihre Präferenzen angeben

Des Weiteren wurde die Bewertungsfunktion gemäß den Nutzerpräferenzen umgesetzt. Das Design orientiert sich hierbei am Prototyp.

d) Handlungsempfehlungen integrieren

Eine neue Funktion, die in das Labormuster integriert wurde, waren die Handlungsempfehlungen. Nachdem Nutzerinnen und Nutzer bestimmte Datenverarbeitungen im App-Kontext bewertet haben und damit anzeigen, dass diese nicht mit ihren Präferenzen übereinstimmen, werden im Labormuster automatisch die Handlungsempfehlungen aus den Informationstexten angezeigt. Ziel ist es hierbei, durch präzise und einfach verständliche Informationen Nutzerinnen und Nutzer zum Handeln zu motivieren.

e) Transparenz bei Datenverarbeitungen des Labormusters

Weitere neue Funktionen des Labormusters finden sich in den Bereichen „Partner“, „Datenschutzerklärung“ und „Impressum“. Als Anwendung, die die Datenverarbeitungen anderer Anbieter auswertet und für bestimmte Datenverarbeitungen „rote Linien“[78] markiert, ist es notwendig, selbst ein hohes Maß an Transparenz in das Labormuster zu integrieren. In den genannten Bereichen wurden deshalb entwurfsweise Informationen angelegt, die über das Projekt, die erhobenen Daten und Verantwortlichkeiten informieren.

Die folgende Abbildung gibt einen Überblick über das Labormuster 2017.

[78] Siehe 4.8.2.2.

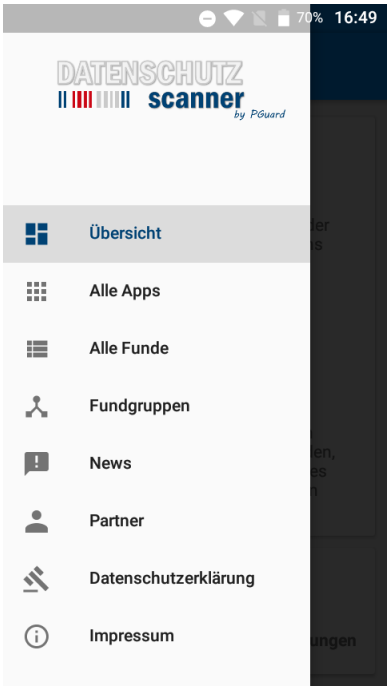


Abbildung 73: Labormuster 2017 – Menüansicht

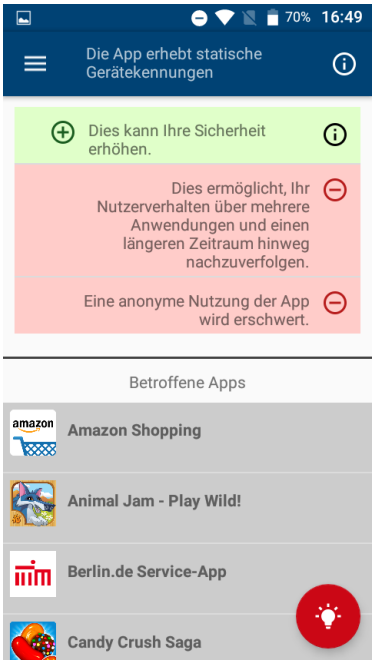


Abbildung 74: Labormuster 2017 – Informationstext der Datenverarbeitungskategorie „Die App erhebt statische Gerätekennungen“[79]

[79] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.

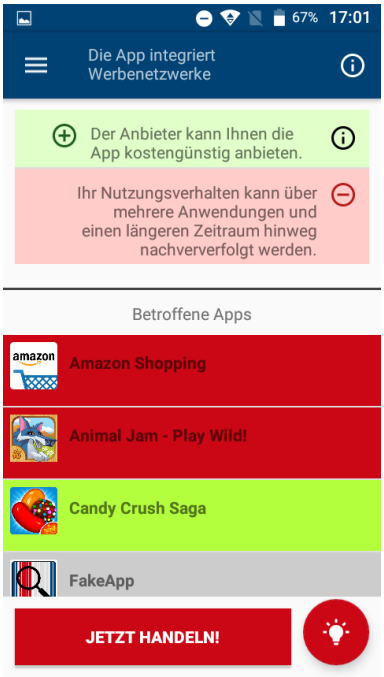


Abbildung 75: Labormuster 2017 – Möglichkeit, eine Handlungsempfehlung anzufordern[80]

5.2.6 User-Tests auf der Internationalen Funkausstellung Berlin

Auf der Internationalen Funkausstellung (IFA), die vom 01.09. bis zum 06.09.2017 in Berlin stattfand, stellte das Konsortium das Labormuster 2017 vor. Da sich die Deutsche Telekom AG für datenschutz- und verbraucherfreundliche Lösungen einsetzt und die App als eine besonders gelungene Umsetzung ansah, bot sie an, eine Präsentation des Labormusters auf dem IFA-Messestand der Deutschen Telekom AG zu ermöglichen. Durch diese Einladung war es möglich, das Labormuster an potentiellen Nutzerinnen und Nutzern zu testen. Ziel der Testung war es, erste Einblicke in Nutzerreaktionen zu gewinnen und die Nutzerführung und die App-Oberfläche auf den Prüfstand zu stellen.

5.2.6.1 Methodik: User-Tests auf der Internationalen Funkausstellung Berlin

Zur Vorbereitung der Nutzergespräche wurden Leitfragen entwickelt, die von den Konsortialmitgliedern, die auf der IFA vertreten waren, an Nutzerinnen und Nutzer gestellt werden sollten. Die Leitfragen enthielten u.a. die folgenden Fragen:

(1) Ist das Konzept des Labormusters verständlich?
(2) Ist das Labormuster interessant und bietet Mehrwert?
(3) Welches generelle Feedback geben Nutzerinnen und Nutzer zum Labormuster?

Das Feedback wurde von den Konsortialmitgliedern während der IFA gesammelt und hiernach aufbereitet. Insgesamt wurden über 80 Gespräche geführt, deren Dauer zwischen 5 und 60 Minuten lag.

[80] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.

5.2.6.2 Ergebnisse: User-Tests auf der Internationalen Funkausstellung Berlin

Insgesamt erhielt das Labormuster eine positive Rückmeldung. In den Gesprächen mit potentiellen Nutzerinnen und Nutzern konnte festgestellt werden, dass die App einen Mehrwert bietet, da sich viele Nutzerinnen und Nutzer derzeit mit dem Schutz ihrer persönlichen Daten im Smartphone überfordert fühlen. Der Aspekt der Verständlichkeit wurde in den Gesprächen häufig hervorgehoben und Gesprächspartnerinnen und -partner wiesen darauf hin, dass sie sich vor allem wünschen, dass komplexe Sachverhalte in der Anwendung verständlich dargestellt werden.

Während einige Gesprächspartnerinnen und -partner es positiv hervorhoben, dass im Labormuster keine Vorbewertungen der Datenverarbeitungen vorgenommen werden, teilten andere mit, dass eine Vorbewertung und Warnung vor negativen Datenverarbeitungen eine hilfreiche Funktion darstellen könnte.

Ein weiterer Aspekt, der mehrheitlich positiv von den Gesprächspartnerinnen und -partnern hervorgehoben wurde, war das Anzeigen der Handlungsempfehlungen im Labormuster. Die potentiellen Nutzerinnen und Nutzer gaben an, dass diese Unterstützung neue und hilfreiche Informationen für sie bieten würde. Zusätzlich wurde jedoch auch der Wunsch geäußert, die Umsetzung der Handlungsempfehlungen einfacher zu gestalten und in die Anwendung zu integrieren.

5.2.7 Experten Beta-Phase

Nach der Testung auf der Internationalen Funkausstellung Berlin wurden Optimierungen der App vorgenommen und ein weiterer User-Test in Form einer Experten Beta-Phase durchgeführt. Ziel der Experten-Beta-Phase war zum einen eine weitere Qualitätskontrolle der Inhalte des Labormusters (App), zum anderen sollte analysiert werden, ob der Mehrwert der App auch von Expertinnen und Experten als positiv bewertet und die App ihrer Mission gerecht wird.

5.2.7.1 Methodik: Experten Beta-Phase

Die Beta-Phase wurde von Dezember 2017 bis Februar 2018 durchgeführt und ausgewertet. Folgende Vorbereitungen wurden hierbei im Einzelnen getroffen:

- Auswahl relevanter Expertinnen und Experten aus Politik, Wirtschaft und Zivilgesellschaft
- Rekrutierung der Expertinnen und Experten
- Erstellung eines Tutorials zur Nutzung des Labormusters[81]
- Technische Bereitstellung des Labormusters für die Expertinnen und Experten
- Einrichtung eines Support-Kanals für etwaige Fragen der Expertinnen und Experten

Zur Erhebung des Expertenfeedbacks wurden Experteninterviews vorbereitet und terminiert. Für die Interviews wurden Leitfragen zu fünf Themenbereichen formuliert (Tabelle 7) und den Expertinnen und Experten im Vorfeld der Interviews zugesendet.

[81] Das verwendete Tutorial ist im Anhang unter 9.5 abgebildet.

Insgesamt wurden sechs Interviews á 30 Minuten mit Vertreterinnen und Vertretern aus Wissenschaft, Unternehmen, Politik und Zivilgesellschaft geführt, transkribiert und ausgewertet.

Themenbereich 1: Allgemeine Bewertung • Wie hat Ihnen die App insgesamt gefallen? Warum? **Themenbereich 2: Inhalte und Funde** • Welche Inhalte (bspw. Funde oder Informationen zur Datenverarbeitung) fanden Sie besonders hilfreich? Welche besonders interessant? Welche weniger interessant? • Wie fanden Sie die inhaltliche Darstellung der Konsequenzen einer Datenverarbeitung, das heißt die Pro- und Contra-Liste, unterhalb der Funde? • Existieren nach Ihrer Meinung Lücken in der Aufbereitung der Pros und Contras? • Können Sie sich Inhalte vorstellen, die Sie oder Ihre Institution der App zur Verfügung (Integration) stellen können bzw. auf die verlinkt / verwiesen werden könnten? **Themenbereich 3: Funktionen** • Welche Funktionen fanden Sie nützlich? Welche nicht? Warum? • Können Sie sich Funktionen vorstellen, um die die App ergänzt werden könnte? **Themenbereich 4: Usability** • Wie bewerten Sie die Nutzerführung der App insgesamt? Was hat Ihnen besonders gefallen? An welchen Stellen hat es gehakt? **Themenbereich 5: Mehrwert der App** • Glauben Sie Verbraucherinnen und Verbraucher würden diese App nutzen? Gibt es bestimmte Verbrauchergruppen, die ein besonders hohes Interesse haben könnten? • Wird die App Ihrer Mission gerecht „Verbraucherinnen und Verbrauchern mehr Transparenz und Kontrolle für den Selbstdatenschutz zu bieten“?

Tabelle 7: Leitfragen im Rahmen der Experteninterviews

5.2.7.2 Ergebnisse: Experten Beta-Phase

5.2.7.2.1 Themenbereich 1: Allgemeine Bewertung

Im Allgemeinen fiel die Bewertung der Expertinnen und Experten positiv aus. So wurde hervorgehoben, dass die Inhalte interessant seien und die App das Potenzial habe, Datenschutzbewusstsein bei Verbraucherinnen und Verbrauchern zu schaffen. Jedoch wurden auch Verbesserungspotentiale durch die Expertinnen und Experten aufgezeigt. Diese lagen besonders im Bereich der technischen Umsetzung der Nutzeroberflächen und -führung.

5.2.7.2.2 Themenbereich 2: Inhalte und Funde

Bei der Frage, welche der angezeigten Inhalte (das heißt Informationstexte) besonders interessant für Verbraucherinnen und Verbraucher seien, ergab sich ein geteiltes Bild. Zum einen wurde von den Ex-

pertinnen und Experten hervorgehoben, dass diejenigen Inhalte bzw. Informationstexte besonders interessant seien, die nicht durch Nutzerinnen und Nutzer beeinflusst werden könnten. Andere Expertinnen und Experten hoben hervor, dass besonders Inhalte bzw. Informationstexte zu den Datenschutzerklärungen interessant seien.

Die Verständlichkeit der Inhalte bzw. Informationstexte wurde im Allgemeinen als sehr gut bewertet und auch das integrierte Konzept für die Inhalte bzw. Informationstexte sowohl die negativen als auch die positiven Konsequenzen anzuzeigen, wurde begrüßt. So betonten die Expertinnen und Experten, dass die Anzeige der positiven und negativen Konsequenzen die Nutzerinnen und Nutzer befähige, selbst über Datenverarbeitungen zu entscheiden.

Bezüglich des Detailgrades der Inhalte bzw. Informationstexte existierten bei den Expertinnen und Experten geteilte Meinungen. Während die Hälfte der Expertinnen und Experten angab, dass die Inhalte bzw. Informationstexte mehr Informationen benötigen, gab die andere Hälfte der Expertinnen und Experten an, dass die Inhalte bzw. Informationstexte, insbesondere die Angaben zu den Pro- und Contras, schon zu ausführlich seien.

Gefragt nach externen Inhalten, die zur besseren Informiertheit der Nutzerinnen und Nutzer in die App integriert werden könnten, nannten die Expertinnen und Experten vor allem die Angebote der Verbraucherzentralen und des BMJV-geförderten Projekts "mobilsicher".[82]

5.2.7.2.3 Themenbereich 3: Funktionen

Eine Funktion, die von den Expertinnen und Experten gewünscht wurde, war die Vorbewertung eindeutiger Gesetzesverstöße. Das heißt, Inhalte bzw. Informationstexte, die ausschließlich negative Konsequenzen beinhalten und einen eindeutigen Gesetzesverstoß darstellen, sollten als „Rote Linien“[83] definiert werden. Die in 9.1 präsentierten finalen Informationstexte sind aus diesem Grund entsprechend markiert.

Darüber hinaus wurden von den Expertinnen und Experten Funktionen genannt, die die Nutzung der App attraktiver machen könnten. Hierzu zählen:

1. Informationen über eine App vor der Installation auf dem Endgerät ermöglichen
2. Hinweise oder Zugang zu Auskunftsverlangen ermöglichen
3. Bei Verstößen in der Datenverarbeitung Kontakt zu den zuständigen Aufsichtsbehörden ermöglichen
4. Berechtigungsmanagement bzw. Einstellung der App-Zugriffe aus der App heraus komfortabel ermöglichen
5. Erweiterung der Bewertungsfunktion bspw. durch Pauschalbewertungen und selbstlernende Bewertung
6. Erstellen eines Nutzerprofils ermöglichen
7. Optimierung der bildlichen Darstellung der Inhalte bzw. Informationstexte, bspw. durch Icons

[82] siehe https://www.surfer-haben-rechte.de oder https://mobilsicher.de
[83] Siehe 4.8.2.2.

5.2.7.2.4 Themenbereich 4: Usability

Die Expertinnen und Experten sahen die größten Schwachstellen der App in ihrer Nutzerführung. Speziell die mangelnde Intuition bei der Bedienung wurde kritisiert und es wurde empfohlen, die App „moderner und ansprechender“ zu gestalten.

5.2.7.2.5 Themenbereich 5: Mehrwert der App

Gefragt nach dem Mehrwert der App ergab sich ein zweigeteiltes Bild. Zum einen wurde die Mission „Transparenz“ zu schaffen positiv hervorgehoben und als erfüllt angesehen. Der Aspekt der Ausübung von „Kontrolle“ wurde nur bedingt als erfüllt angesehen. Begründet wurde diese Bewertung vor allem durch die der nicht vollständig intuitiven Bedienung und der fehlenden Umsetzungsmöglichkeiten von Handlungsempfehlungen aus der App heraus.

5.2.8 Labormuster 2018: Version des weiterentwickelten Prototyps

Für das Labormuster 2018 wurden die Funktionen aus dem Labormuster 2017 übernommen. Die Nutzerführung wurde zur besseren Bedienbarkeit kondensiert und die visuelle Darstellung ansprechender gestaltet. Außerdem wurden in das Labortmuster 2018 die finalen Informationstexte[84] eingefügt.

Eine weitere Funktion wurde für das Labormuster 2018 überarbeitet: Transparenz bei Datenverarbeitungen des Labormusters.

Die Datenschutzerklärung des Labormusters wurde gemäß der zum 25. Mai 2018 in Kraft getretenen Datenschutzgrundverordnung (DSGVO) angepasst und in das Labormuster integriert.

Die folgenden Abbildungen zeigen Ausschnitte aus dem aktuellen Labormuster:

[84] Siehe 9.1.

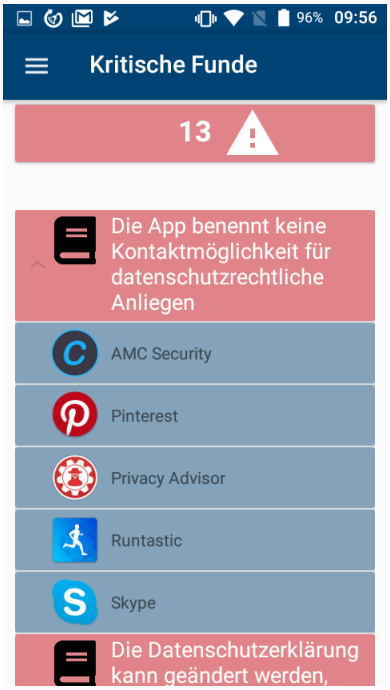


Abbildung 76: Labormuster 2018 – Anzeige der kritischen Funde und die jeweils betroffenen Apps[85]

Abbildung 77: Labormuster 2018 – Anzeige eines Fundes, dessen Informationstext und die jeweils betroffenen Apps mit der direkten Möglichkeit zur selbständigen Bewertung[86]

[85] Die verwendeten Beispiele und App-Logos innerhalb der Abbildung dienen lediglich der Illustration des Nutzerführungskonzepts und sagen nichts über tatsächliche Datenverarbeitungen der Anbieter aus.
[86] Die verwendeten Beispiele und App-Logos innerhalb der Abbildung dienen lediglich der Illustration des Nutzerführungskonzepts und sagen nichts über tatsächliche Datenverarbeitungen der Anbieter aus.

Abbildung 78: Labormuster 2018 – Handlungsempfehlung und Möglichkeit, dieser Empfehlung umzusetzen[87]

Im Rahmen des Forschungsprojektes wurde auch eine Vielzahl weiterer Funktionen angedacht, konzipiert und deren Umsetzbarkeiten erforscht. Die konkrete Implementierung wurde aber nicht immer vorgenommen, da diese Funktionen im Rahmen des Forschungsprojektes nicht mit der hinreichenden Qualität – sowohl hinsichtlich Verlässlichkeit als auch Nutzerfreundlichkeit – umgesetzt werden konnten.

Es wurde sich stattdessen dafür entschieden, den konkret implementierten Funktionsumfang zu fokussieren und in seiner Qualität einer dem Forschungsziel angemessenen Umsetzung zuzuführen. Eine rein quantitative Anhäufung von Funktionen hätte dem Forschungsziel entgegengestanden.

Dennoch soll im Rahmen dieses Berichts auf angedachte Funktionen und deren Mehrwerte für Nutzerinnen und Nutzer beispielhaft eingegangen werden.

5.2.8.1 Analyse hinsichtlich abweichender Datenschutzerklärungen

Bei der manuellen Aufbereitung der Datenschutzerklärungen war festzustellen, dass Apps zwar eine Datenschutzerklärung im Google Play Store verlinken, diese aber nicht mit der Datenschutzerklärungen übereinstimmt, die in der App angezeigt wird. Die Abweichungen hatten unterschiedlichen Hintergründe:

- die im Google Play Store verlinkte Datenschutzerklärung war lediglich eine generische Datenschutzerklärung z.B. für die Webseite des Anbieters, ohne die App zu berücksichtigen

[87] Die verwendeten Beispiele und App-Logos innerhalb der Abbildung dienen lediglich der Illustration des Nutzerführungskonzepts und sagen nichts über tatsächliche Datenverarbeitungen der Anbieter aus.

- die im Google Play Store verlinkte Datenschutzerklärung und die in der App bereitgestellte – App-spezifische – Datenschutzerklärung waren auf einem abweichenden Stand

Die Auswirkungen für Nutzerinnen und Nutzer können im Einzelfall erheblich sein. Insofern wurde es für sachdienlichen erachtet, Nutzerinnen und Nutzern sowohl einen Hinweis auf diese Abweichungen zu geben als auch die konkreten Abweichungen optisch aufbereitet nachvollziehen zu können. Auf lange Sicht wäre es natürlich wünschenswert, die Abweichungen hinsichtlich der datenschutzrechtlichen Relevanz automatisiert vorbewerten zu können.

Die Technik für eine optische Darstellung etwaiger Abweichungen existiert, siehe 4.10.12. Sie müsste jedoch für die Darstellungen auf kleineren, mobilen Endgeräten optimiert werden. Grundsätzlich ist es technisch auch möglich, Datenschutzerklärungen aus den Apps auszulesen, 4.10.5. Datenschutzerklärungen werden innerhalb von Apps jedoch häufig an unterschiedlichen Stellen bereitgestellt, sodass automatisierte Crawler keine hinreichend verlässlichen Informationen bereitstellen können. Insbesondere, da die Auswertung der auf diesem Weg gefundenen Datenschutzerklärungen wiederum auf einer OCR[88]-Aufbereitung basiert, die selbst Ungenauigkeiten auslösen kann.

Ebenso ist das Auslesen der in einer App bereitgestellten Datenschutzerklärung auf Code-Ebene als nicht hinreichend zuverlässig einzustufen, da sowohl Obfuskationsmethoden angewendet werden können als auch die Integration durch Nachladen des konkreten Textes von einer externen Quelle erfolgen kann, mithin die Datenschutzerklärung nicht innerhalb der App codiert wäre.

5.2.8.2 Unzutreffende Angaben in einer Datenschutzerklärung

Ziel des Forschungsprojektes war es insbesondere auch Abweichungen zwischen tatsächlichen Datenverarbeitungen und den bereitgestellten Datenschutzerklärungen herauszuarbeiten.

Die für eine solche Aussage erforderlichen Logiken und Analysetechniken konnten im Wesentlichen entwickelt werden, nähere Ausführungen in den 4.9 und 4.10. Für den Demonstrator wurde sich dennoch gegen eine Implementierung entschieden. Hintergrund sind häufig nicht App-spezifische Datenschutzerklärungen, die die Datenverarbeitungen mehrerer Dienste zusammenfassen. Hierdurch ist es möglich, dass die automatisierte Analyse eine Nichtverarbeitung in einem anderen, nicht App-bezogenen Kontext feststellt, obgleich eine Verarbeitung für die App tatsächlich nicht ausgeschlossen wurde. Gleiches gilt für Ausschlüsse ausschließlich für dezidierte Verarbeitungszwecke. Derart kontextualisierte Analysen der Datenschutzerklärung lieferten nicht die hinreichend zuverlässigen Ergebnisse und waren für einen Großteil der sachdienlichen Informationen für Nutzerinnen und Nutzer auch nicht erforderlich.

Zudem galt es zu berücksichtigen, dass hinreichend nachvollziehbare und reproduzierbare Einzelergebnisse der semantischen und technischen Analyse genügen, um eine Aussage dahingehend zu treffen, ob eine Datenverarbeitung stattfindet (stattfinden könnte). In einer abhängigen Kombination können

[88] Optical Character Recognition –automatisierte Texterkennung.

sich die individuellen Toleranzen indessen potenzieren, welches sodann zu nicht mehr nachvollziehbaren und belastbaren Ergebnissen führen kann.

Nicht zuletzt ist anzumerken, dass das Datenschutzrecht ohnehin als Verbot mit Erlaubnisvorbehalt ausgestaltet ist. Datenschutzerklärungen sind in Ihrer Transparenzfunktion somit als „Positiv“ konzipiert, nicht jedoch als Negativ. Explizite Formulierungen, die eine Datenverarbeitung ausschließen, konnten somit äußerst selten festgestellt werden. Daten, die nicht als in einer Datenschutzerklärung benannt sind, könnten allerdings als „ausgeschlossen“ verstanden werden. Eine solche automatisierte Logik kollidiert jedoch mit – juristisch nicht per se unzulässigen – Ungenauigkeiten und Beispielaufzählungen in Datenschutzerklärungen. Mithin ist es möglich, dass die konkreten Daten in einer Datenschutzerklärungen nicht positiv aufgezählt sind, also auch nicht im Rahmen einer automatischen Analyse gefunden werden können, obgleich diese aus juristischer Sicht sehr wohl vom angegebenen Verarbeitungsumfang abgebildet sein könnten.

Zwar hätte auch trotz dieser Umstände eine Information im Rahmen des Demonstrators implementiert werden können. Sodann hätten die vorangestellten Herausforderungen und eventuellen Ungenauigkeiten den Nutzerinnen und Nutzern ebenfalls verständlich kommuniziert werden müssen. Auf dieser Ebene erschien keine sinnvolle Darstellungsoption denkbar, die den Nutzerinnen und Nutzern einen tatsächlichen Mehrwert bietet.

5.2.9 Zusammenfassung

Allgemein können sowohl das Feedback im Rahmen der Internationalen Funkausstellung Berlin als auch der Experten Beta-Phase als Erfolg gewertet werden. Die inhaltlichen Kernbestandteile der App wie beispielweise die Informationstexte wurden grundsätzlich als nützlich bewertet und können Verbraucherinnen und Verbrauchern einen großen Mehrwert bieten.

Im Rahmen technischer Verbesserungspotentiale ist an eine automatisierte Vorbewertung von Abweichungen der Datenschutzerklärung in der App und im Google Play Store (derzeit optisch darstellbar) zu denken, sowie an eine technische Implementierung der Möglichkeit, tatsächliche technische Verarbeitungen und die in der Datenschutzerklärung angegebene Verarbeitung zu vergleichen.

5.3 Technische Anforderungen an den App-Client

5.3.1 Privacy-by-design und Pseudonymisierung des App-Clients

Zur Übermittlung und Verarbeitung der vom DATENSCHUTZscanner übermittelten Daten wurde eine pseudonyme Zuordnung integriert, deren Hintergründe und Funktionsweise im Folgenden näher beschrieben werden.

Konzeptionell sollte eine Zuweisung von Anfragen des Clients zu später gelieferten Informationen möglich sein. Allerdings sollte dabei kein Login der Betroffenen erforderlich sein. Dennoch sollte eine Trennung der Informationen zwischen den unterschiedlichen anfragenden Instanzen realisiert werden, ebenso sollte es möglich sein, statistisch auszuwerten, wie viele Clients installiert sind bzw. aktiv genutzt werden.

Hierzu wurde eine Trennung von zwei pseudonymen IDs vorgesehen, wovon eine beim ersten Start des App-Clients zufällig generiert wird und die andere je Anfrage zu noch nicht verfügbaren Testergebnissen. Technisch konnte dieser Ansatz erfolgreich implementiert und verifiziert werden. Allerdings bedurfte dieser Ansatz sowohl innerhalb der Clients bzw. auf dem Endgerät der Betroffenen als auch auf dem Server respektive innerhalb des PGuard-Backends viele Ressourcen. Dieser Ressourcenbedarf galt es im Weiteren zu reduzieren.

Client-Anfragen wurden durch eine bessere Aufbereitung so optimiert, dass diese wesentlich effizienter und letztlich schneller abgearbeitet werden konnten. Hierdurch entfiel ein Logging im PGuard-Backend. Es werden auf Wunsch der Betroffenen lediglich die für konkret relevanten Testergebnisse angefragt: soweit bereits Testergebnisse vorliegen, werden diese unmittelbar an den App-Client übermittelt. Soweit noch Testergebnisse vorliegen, werden entsprechende Ergebnisse der betroffenen Apps durch das PGuard-Backend bei den beiden Analyse-Backends (technisch und semantisch) angefragt. Sobald der App-Client erneut Testergebnisse anfragt, werden die in der Zwischenzeit von den Analyse-Backends an das PGuard-Backend übermittelten Testergebnisse auch an den App-Client übertragen, ohne dass eine Verknüpfung zwischen den beiden Anfragen erstellt wird oder erstellt werden kann. Lediglich die per Opt-In erhobenen zufälligen IDs, die sich auch einfach durch das Löschen des Zwischenspeichers oder durch De- und wieder Reaktivieren der Opt-In Funktion neu iterieren lassen, werden übermittelt, um die Anzahl von DATENSCHUTZscanner Installationen einschätzen zu können. Diese wesentlich schlankere Realisierung erlaubte es jegliches Tracking auf ein optionales Minimum zu reduzieren und daher dem Anspruch des „privacy-by-design“ gerecht zu werden.

5.3.2 Besonderheiten bei der Bereitstellung relevanter Features für Nutzerinnen und Nutzer

5.3.2.1 Android

Bei Android ergab sich eine unerwartete Hürde bei der Integration der Nutzungsstatistiken zu den installierten Applikationen. Da bei der Entscheidungsfindung der notwendigen Maßnahmen bezüglich einer problematischen Applikation zum Beispiel das Datum der letzten Nutzung für Betroffene Einfluss auf die Notwendigkeit einer Handlung haben könnte, sollten diese Statistiken in der Detailansicht der jeweiligen App übersichtlich dargestellt werden. Android stellt dafür den „UsageStatsManager“[89] zur Verfügung, der jedoch im entwickelten Labormuster keine zuverlässigen Daten zur Verfügung stellte. Die dargestellten Informationen und deren Fehlerhaftigkeit unterschied sich auch je Endgerät; so waren die Informationen teils veraltet, teils gar nicht verfügbar. Vor diesem Hintergrund wäre ein solches Feature für Nutzerinnen und Nutzer nahezu nutzlos. Die Funktion und die technische Implementierung wurde dennoch im Demonstrator belassen, da sich eine fehlerfreie Funktionalität zu einem späteren Zeitpunkt, etwa durch ein Android-Update oder eine alternative Informationsquelle, stabilere Daten hätte ergeben können.

[89] https://developer.android.com/reference/android/app/usage/UsageStatsManager

Abbildung 79 Letztes Nutzungsdatum in der Detailansicht

5.3.2.2 iOS

im Gegensatz zu Android können Applikationen unter iOS andere auf dem Gerät installierten Anwendungen nicht ohne Weiteres identifizieren. Bis iOS 10 war dies zwar über inoffizielle APIs möglich, jedoch würden Applikationen, die auf diese APIs zurückgreifen, inzwischen im Rahmen einer Revision abgelehnt und / oder aus dem App Store entfernt. Es bestehen allerdings auch unter iOS technische Möglichkeiten, wie Nutzerinnen und Nutzern ähnliche – nutzerfreundliche – Funktionen bereitgestellt werden können. Zum Beispiel konnte für das Forschungsvorhaben eine Lösung erarbeitet werden, bei welcher mittels Mobile-Device-Management („MDM“) das App-Porfolio des Gerätes über das MDM ausgelesen und sodann vom Server des MDM wieder an die Applikation auf dem Endgerät zurückübertragen werden können. Dadurch lässt sich das auf dem Gerät eingesetzte Applikationsportfolio darstellen und dementsprechend eine Handlungsempfehlung aussprechen.

Derartige Ansätze sind in B2B-Anwendungen bereits implementiert, sodass von einer technischen Realisierbarkeit auszugehen ist. Allerdings nutzen Privatpersonen üblicherweise kein MDM, sodass im Rahmen des Forschungsprojektes dieser Lösungsansatz nicht im Detail weiterverfolgt wurde.

Ebenfalls wäre es unter iOS denkbar, dass Nutzerinnen und Nutzer die auf dem Endgerät installierten Applikationen selbst und manuell dem App-Client mitteilen. Eine derartige Lösung ist aber bezüglich der Nutzerfreundlichkeit erheblich schlechter und könnte im Ergebnis für Nutzerinnen und Nutzer eine zu große Hürde darstellen. Insbesondere ist zu beachten, dass die bloße Angabe der installierten App unter Umständen nicht ausreicht, um hinreichend präzise Ergebnisse und Handlungsempfehlungen zu erhalten; vielmehr erscheint auch die konkret installierte App-Version relevant, die die wenigsten Nutzerinnern und Nutzer kennen. In diesen Fällen müssten die Nutzerinnen und Nutzer also zudem die konkrete App-Version zuvor auf dem Endgerät auslesen und – bei jedem Update einer installierten App – ebenfalls manuell anpassen.

5.3.3 Zusammenfassung

Für die Ausgestaltung der technischen Rahmenbedingungen wurde darauf geachtet, dass der App-Client möglichst datenschutzfreundlich ausgestaltet wird. Zudem galt es für den App-Client und die darin integrierten Funktionen Umsetzungsmöglichkeiten zu finden. Hierbei konnten Unterschiede zwischen den Betriebssystemen festgestellt werden.

5.4 Datenschutzrechtliche Anforderungen an die durch einen App-Client bereitgestellten Funktionen

Bei der Bereitstellung eines App-Clients sind stets datenschutzrechtliche Anforderungen zu beachten. Für das Forschungsvorhaben waren hierbei mehrere Aspekte zu berücksichtigen, die neben den allgemeinen Anforderungen – Erstellung einer Datenschutzerklärung, Abklärung der rechtlichen Erlaubnistatbestände für mögliche Datenverarbeitungen – mit spezifischen Herausforderungen einhergingen.

Für die Glaubwürdigkeit des Forschungsvorhabens war es von besonderer Bedeutung, dass im Bereich der datenschutzrechtlichen Anforderungen und deren technischer Umsetzung diese im Sinne des Grundsatzes privacy-by-design Berücksichtigung fanden. Ebenso sollte es Betroffenen transparent und nutzerfreundlich ermöglicht werden, datenschutzrechtliche Implikationen in der App selbst konfigurieren zu können.

Zu den einzelnen Aspekten wurden Konzepte entwickelt, die zum Teil in den Demonstrator implementiert wurden, zum Teil im Sinne einer Kosten-Nutzen-Abwägung im Rahmen der Entwicklung keine Berücksichtigung fanden, jedoch als theoretische Umsetzungsoptionen durchdacht wurden.

Datenschutzrechtlich ist jede Information, die einer natürlichen Person zugeordnet werden kann, ein personenbezogenes Datum. Hierbei ist es ausreichend, dass ein Personenbezug nicht direkt existiert, sondern gegebenenfalls hergestellt werden kann.[90]

5.4.1 Auslesen der installierten Apps auf dem Endgerät der Betroffenen

Der Umstand, ob Betroffene eine App installiert haben oder nicht, ist ein personenbezogenes Datum, soweit dieser Information ein Personenbezug zugeordnet werden kann (siehe auch 5.4, 5.4.3.2). Hierbei ist auch zu berücksichtigen, dass dieses Datum für Betroffene eine erhöhte Sensibilität aufweisen kann, soweit sich aus der App-Bezeichnung oder dem Funktionsumfang weitere Rückschlüsse ziehen lassen. Rückschlüsse können etwaig auf den Gesundheitszustand, sexuelle, weltanschauliche oder politische Präferenzen gezogen werden.[91]

Im Sinne des privacy-by-design galt es somit zu klären, ob die Verarbeitung dieser Informationen durch den App-Client erforderlich ist. Datenschutzrechtlich galt es zudem abzuklären, ob die Verarbeitung durch den App-Client bereits eine Verarbeitung durch die verantwortliche Stelle darstellt.

[90] Artikel 4 Ziffer 1 DS-GVO.
[91] So zum Beispiel Dating-Apps mit Fokus auf bestimmte Vorlieben – KNKI: BDSM Dating (https://play.google.com/store/apps/details?id=com.kinkygal.kink) –, Apps von politischen Parteien, Gesundheitsapps zur Dokumentation oder Ermittlung von Erkrankungen – Molexplore – Melanoma & Skin Cancer App (https://play.google.com/store/apps/details?id=com.borealos.medical.molexplore) –, Apps einzelner Kirchen(-gemeinden) und Weltanschauungen.

5.4.1.1 Verarbeitung durch App-Client als Verarbeitung durch die verantwortliche Stelle

Verarbeitung meint „jeden mit oder ohne Hilfe automatisierter Verfahren ausgeführten Vorgang oder jede solche Vorgangsreihe im Zusammenhang mit personenbezogenen Daten wie das Erheben, das Erfassen, die Organisation, das Ordnen, die Speicherung, die Anpassung oder Veränderung, das Auslesen, das Abfragen, die Verwendung, die Offenlegung durch Übermittlung, Verbreitung oder eine andere Form der Bereitstellung, den Abgleich oder die Verknüpfung, die Einschränkung, das Löschen oder die Vernichtung“, Artikel 4 Ziffer 2 DSGVO.

Verantwortlicher ist gemäß Artikel 4 Ziffer 7 DSGVO „die natürliche oder juristische Person, Behörde, Einrichtung oder andere Stelle, die allein oder gemeinsam mit anderen über die Zwecke und Mittel der Verarbeitung von personenbezogenen Daten entscheidet; sind die Zwecke und Mittel dieser Verarbeitung durch das Unionsrecht oder das Recht der Mitgliedstaaten vorgegeben, so kann der Verantwortliche beziehungsweise können die bestimmten Kriterien seiner Benennung nach dem Unionsrecht oder dem Recht der Mitgliedstaaten vorgesehen werden“.

Somit stellt sich die Frage, ob ein Informationszugriff ausschließlich durch einen App-Client – ohne Übermittlung dieser Information an die verantwortliche Stelle – bereits eine Verarbeitung im Sinne des Datenschutzrechts darstellt. Hierzu müsste es zu irgendeinem Zeitpunkt jedenfalls zu einer Verarbeitung in Form einer Erhebung kommen.[92] Die DSGVO liefert – anders als noch das Bundesdatenschutzgesetz[93] – keine Legaldefinition der Erhebung.[94]

Aus der Kombination der Definition des Verantwortlichen und der Definition der Verarbeitung ergibt sich aber, dass eine datenschutzrechtliche Relevanz voraussetzt, dass die verantwortliche Stelle irgendwie gearteten, finalen Einfluss auf die Verarbeitung ausüben kann. Solange die Informationen auf dem Endgerät verbleiben und der Verantwortliche keinerlei Zugriff auf diese Informationen erhält, wäre also mangels Erhebung auch von keiner Verarbeitung im datenschutzrechtlichen Sinne auszugehen.

Alternativ kann jedoch auch die Auffassung vertreten werden, dass die Daten – wenn auch nur lokal – in die technische Verarbeitungsumgebung der App überführt wurden. Der Betroffene selbst hat – in der Regel – keine Möglichkeiten die Verarbeitung durch die App zu beeinflussen. Die verantwortliche Stelle der App – der Anbieter bzw. dessen Entwickler – haben die Hoheit über die Ausgestaltung des App-Quellcodes und die damit einhergehenden Verarbeitungsvorgänge. Obgleich die Daten also nur lokal – und somit auf den ersten Blick außerhalb der Kenntnis und außerhalb des Verarbeitungszugriffs der verantwortlichen Stelle liegen – sind die Informationen möglicherweise bereits in die Sphäre der verantwortlichen Stelle übergegangen.

Im Bereich der Überwachung mit optisch-elektronischen Einrichtungen ist es zudem nach herrschender Meinung irrelevant, ob eine tatsächliche Überwachung stattfindet oder lediglich eine Attrappe installiert wird. Hintergrund ist auch hier, dass das Ob und das Wie letztlich in der ausschließlichen Einflusssphäre

[92] Die Erhebung als Beginn eines jeden Verarbeitungsprozesses; vgl. auch *Bäcker* in Kühling/Buchner, Datenschutz-Grundverordnung, 1. Auflage 2017, Art. 13, Rn. 12.
[93] § 3 Absatz 3 BDSG a.F.
[94] So auch zum Beispiel *Bäcker* in Kühling/Buchner, Datenschutz-Grundverordnung, 1. Auflage 2017, Art. 13, Rn. 12.

der verantwortlichen Stelle liegt. Hierbei ist indessen zu differenzieren zwischen der unmittelbar datenschutzrechtlichen Dimension nach § 6b BDSG a.F. und der zivilrechtlichen Dimension nach §§ 823 Abs. 1, 1004 Abs. 1 BGB, Art. 2 i.V.m. Art. 1 Abs. 1 GG.[95] Scheint der zivilrechtliche Beseitigungs- und Unterlassungsanspruch gefestigt[96], so ist die datenschutzrechtliche Anwendbarkeit von § 6b BDSG eher umstritten. So heißt es in der Orientierungshilfe der Aufsichtsbehörden, dass über die individuelle Auffassung der jeweiligen Behörde hinsichtlich der sachlichen Anwendbarkeit Betroffene diese „auf Anfrage“ erfahren.[97]

Berücksichtigt man die Auffassung zu optisch-elektronischen Einrichtungen zusammen mit der durch die Datenschutzgrundverordnung intendierten, gesteigerten Transparenz für Betroffene, so ist es juristisch argumentierbar, auch bei einer rein lokalen Informationsverarbeitung von einer Verarbeitung im datenschutzrechtlichen Sinne auszugehen. Es ist aber ebenfalls vertretbar, weiterhin eine unterschiedliche Bewertung, je nach Dimension, vorzunehmen. In diesem Falle würde ggf. ein zivilrechtlicher Anspruch aus dem allgemeinen Persönlichkeitsrecht bestehen, das formale Datenschutzrecht aber noch nicht anwendbar sein.

Die praktischen Auswirkungen eines Streitentscheids sind erheblich. Ziel des Forschungsprojekts ist es, die Transparenz und die Selbstbestimmung für Betroffene zu erhöhen. Im Rahmen des Forschungsprojektes konnte ein Streitentscheid unterbleiben, da zumindest der Personenbezug der Daten aus Sicht der verantwortlichen Stelle im Rahmen des Im Forschungsprojekt entwickelten Labormusters vermieden werden konnte, siehe 5.4.3 und 5.4.5.

5.4.1.2 Erforderlichkeit der Verarbeitung der Information durch den App-Client

Die Verarbeitung der Information „installierte Apps auf dem Endgerät des Betroffenen“ ist für den grundsätzlichen Betrieb des App-Clients nicht erforderlich. Für die Bereitstellung der aus Betroffenensicht förderlichen Funktionen mit einem angemessenen Grad an Nutzerfreundlichkeit kann jedoch nicht gänzlich auf diese Information verzichtet werden.

Technisch bestünde die Möglichkeit, die installierten Apps durch die Betroffenen manuell eingeben zu lassen. Hierbei müssten die Betroffenen jedoch den sogenannten Bundle-Identifier ihrer installierten Apps kennen. Betroffene haben eine Vielzahl von Apps auf ihrem mobilen Endgerät installiert. Der Aufwand für Betroffene wäre bei einer manuellen Eingabe erheblich und würde eine aus Sicht der Betroffenen unüberwindbare Hürde darstellen.

Alternativ könnte eine App-Suche entsprechend eines Appstores in den App-Client integriert werden. Hierbei müssten Betroffene lediglich den Namen der App suchen und diese auswählen. Die Integration einer solchen Store-Lösung ist jedoch mit erheblichem Aufwand verbunden, da auf dem Endgerät eine Implementierung der Store-API und der API des PGuard-Backends harmonisiert abgefragt werden müssen, was besonders bei noch nicht geprüften Apps eine spätere Kommunikation der Testergebnisse

[95] vgl. auch Orientierungshilfe des Düsseldorfer Kreises, „Videoüberwachung durch nicht-öffentliche Stellen“, 2014, Seite 5f.
[96] OLG Köln, Urteil vom 30.10.2008 – 21 U 22/08; AG Frankfurt, Urteil vom 14.01.2015 – 33 C 3407/14; AG Lichtenberg, Urteil vom 24.01.2008 – 10 C 156/07; AG Aachen, Urteil vom 11.11.2003 – 10 C 386/03; LG Braunschweig, Urteil vom 18.03.1998 – 12 S 23/97; LG Bonn, Urteil vom 16.11.2004 – 8 S 139/04.
[97] Orientierungshilfe des Düsseldorfer Kreises, „Videoüberwachung durch nicht-öffentliche Stellen“, 2014, Seite 5f.

erschwert. Hinsichtlich des im Rahmen des Forschungsprojektes entwickelten Labormusters und dessen Fokus auf Android bedeutet dies, mangels offizieller API für den Google Play Store die für eine Verwendung zur Verfügung stehenden APIs eine nicht vollständig und offiziell dokumentierte Google-Play-Service Kommunikation nachgebaut haben. Hiermit gehen regelmäßig Einbußen hinsichtlich der Stabilität einher. Auch IT-sicherheitstechnische Aspekte sollten nicht außer Acht gelassen werden. Mithin sollten derartige Lösungen nicht in App-Client verwendet werden, insbesondere wenn und soweit sich diese an Verbraucherinnen und Verbraucher richten.

Für das Forschungsvorhaben konkrete Probleme könnten sich außerdem daraus ergeben, dass es einen Unterschied zwischen der für Betroffene sichtbaren Versionsnummer und sogenannten Buildnummer zu geben scheint.

Das automatische Auslesen der installierten Apps wurde vor diesem Hintergrund und unter “usability” Gesichtspunkten als sachdienlich erachtet. Im Sinne des privacy-by-designs galt es jedoch, die datenschutzrechtlichen Implikationen für Betroffene zu minimieren. Grundsätzlich konnte der App-Client technisch so ausgestaltet werden, dass kein Personenbezug der Informationen für die verantwortliche Stelle hergestellt werden kann.[98] Zudem findet die Datenverarbeitung im Wesentlichen lokal statt. Dennoch wurden in den Demonstrator weitere Maßnahmen integriert, die auch eine allenfalls empfundene datenschutzrechtliche Implikation für Betroffene minimiert.

So wurde in der App ein Pseudo-Zugriffsrecht integriert. Grundsätzlich können Apps unter Android die installierten Apps auf dem Endgerät auslesen. Eine besondere Funktionszugriffsgruppe, die durch Betroffene freigegeben oder entzogen werden kann, ist nicht vorgesehen. Allerdings wurde der Demonstrator so ausgestaltet, dass der App-Client die auf dem Endgerät installierten Apps nur auslesen kann, wenn Betroffene der App dieses Pseudo-Zugriffsrecht gewährt haben. Betroffene sind ohne Gewähr dieses Pseudo-Zugriffsrecht lediglich in der Lage einen „Demo“-Bereich im Demonstrator zu nutzen, in dem die Prüfergebnisse unter Android Nutzerinnen und Nutzern häufig genutzter Apps angezeigt werden.

Zudem integriert der Demonstrator für Betroffene die Möglichkeit, einzelne Apps von der Übertragung an das zentrale Prüflabor zum Abruf der Prüfergebnisse auszuschließen.

5.4.2 Technische Analyse – auf dem Endgerät vs. Im zentralen Prüflabor

Die durch den App-Client bereitgestellten Informationen basieren auf zwei wesentlichen Datenquellen, einer technischen und einer semantischen Analyse.

Die semantische Analyse gibt Auskunft über die Inhalte der Datenschutzerklärung einer App. Hier greift diese Analyse auf im Forschungsprojekt erstellte Referenzdaten zurück, sodass diese Analyse zentral stattfindet.

[98] Siehe 5.3.1.

Die technische Analyse nutzt als Datenquelle die zu analysierende App selbst und gibt Aufschluss über die technischen Vorgänge der App. Eine solche Analyse kann grundsätzlich zentral oder auf einem Endgerät stattfinden. Beide Optionen haben bezüglich der Ergebnisqualität Vor- und Nachteile. Nähere Informationen können unter 5.3 und 5.4 nachgelesen werden.

Eine Analyse auf dem Endgerät geht indessen mit dem Nachteil einher, dass der App-Client für vergleichbare Ergebnisse den Datenfluss der Betroffenen mitschneiden und analysieren müsste. Dies hätte eine besonders kritische datenschutzrechtliche Implikation, da die Art der im Datenfluss vorkommenden Daten vorher nicht abzuschätzen ist. Weiter wäre ein solcher Eingriff in das Telekommunikationsgeheimnis des Betroffenen nur schwer zu rechtfertigen und würde höchstwahrscheinlich auch die Daten Dritter betreffen, die in diese Analyse mangels Kenntnis nicht einwilligen könnten. Im Rahmen des Forschungsprojektes wurde sich somit für eine Analyse im zentralen Prüflabor entschieden.

5.4.3 Übertragen der installierten Apps an das Prüflabor

Entsprechend der Entscheidung unter 5.4.2 findet eine Analyse zentral in einem Prüflabor statt. Mithin galt es zu klären, wie Betroffenen die Prüfergebnisse zu den auf dem individuellen Endgerät installierten Apps zugänglich gemacht werden können.

Datenschutzrechtlich waren auch hier wieder unterschiedliche Aspekte zu betrachten. Ist es notwendig, installierte Apps an die verantwortliche Stelle zu übermitteln? Stellen „installierte Apps“ ein personenbezogenes Datum dar? Kann der Personenbezug vermieden werden?

5.4.3.1 Erforderlichkeit der Übermittlung installierter Apps

Betroffene müssen die Prüfergebnisse erhalten, damit diese auf dem lokalen App-Client angezeigt werden können. Es ist technisch möglich, dass die gesamten Prüfergebnisse zu allen Apps und deren Versionen, in den lokalen App-Client integriert werden. In diesem Fall würde die „Anzeige- und Filterlogik“ lokal bei den Betroffenen stattfinden. Eine Übermittlung der installierten Apps an das zentrale Prüflabor wäre nicht erforderlich.

Dies geht allerdings mit technischen Nachteilen für Betroffene einher, unter anderem ein auf Dauer erheblichen Speicherbedarf, eine kontinuierliche starke Datenvolumennutzung und ein verstärkter Stromverbrauch des App-Clients auf dem Endgerät der Betroffenen.

Zudem erfordert dies ein ständiges Prüfen und Analysieren aller Apps in den Appstores, was mit einem hohen, unverhältnismäßigen Ressourceneinsatz der Prüftechnologie einherginge, denn es müssten alle Versionen aller Apps – am besten auch in allen Stores –getestet werden.

Es könnte zur Optimierung des Ressourceneinsatzes eine Fokussierung auf bestimmte „relevante“ Apps erfolgen. Dies ginge jedoch mit weiteren Nachteilen einher. Erstens ist der Begriff „relevant“ äußerst schwammig und kann sich je nach Zielgruppe erheblich unterscheiden. Es bleibt also entweder bei einem großen Anteil von Apps, die durch Betroffene nie genutzt werden, oder aber die standardisierte Prüfliste weist Lücken auf, sodass Betroffene zu den lokal installierten Apps keine Prüfergebnisse erhalten werden. Zielführender hinsichtlich all dieser Aspekte ist die Übermittlung der installierten Apps an das Prüflabor. Hierdurch bedarf der App-Client nur geringen Speicherplatz. Durch die somit technisch

mögliche Anfrage konkreter Prüfergebnisse kann zudem der Speicherplatzbedarf auch auf Dauer geringgehalten werden, da keine unnötigen Prüfergebnisse auf dem lokalen Endgerät gespeichert werden müssen.

Auch die Prüftechnologie kann hinsichtlich der aufgewendeten Ressourcen optimal eingesetzt werden. Es können standardisiert die Apps vorgeprüft werden, bei denen entsprechend der Download-Zahlen in den Appstores eine hohe Verbreitung zu erwarten ist. Die Apps, die hierbei nicht geprüft werden, aber bei den Betroffenen dennoch installiert sind, können nach Erhalt der Anfrage eines Betroffenen in die Prüftechnologie überführt werden, sodass Betroffene beim nächsten Abruf der Prüfergebnisse bereits die sodann neu generierten Prüfergebnisse erhalten können, hierzu auch weitere Informationen unter 4.11.1.2.

5.4.3.2 Sind „installierte Apps“ ein personenbezogenes Datum?

Wie unter 5.4.1 bereits erwähnt, stellen auch installierte Apps ein personenbezogenes Datum dar, solange und soweit diese Information sich auf eine identifizierte oder identifizierbare natürliche Person bezieht. Hierbei kann, wie unter 5.4.1 ausgeführt, die Information unter Umständen diejenige Information, welche eine besondere Kategorie personenbezogener Daten im Sinne des Artikel 9 DSGVO darstellt, (mit-)umfassen.

5.4.3.3 Besteht ein Personenbezug?

Fraglich ist, ob Betroffene eine identifizierte oder identifizierbare Person darstellen, mithin also ein Personenbezug vorliegt.

Als identifizierbar wird eine natürliche Person angesehen, die direkt oder indirekt, insbesondere mittels Zuordnung zu einer Kennung wie einem Namen, zu einer Kennnummer, zu Standortdaten, zu einer Online-Kennung oder zu einem oder mehreren besonderen Merkmalen, die Ausdruck der physischen, physiologischen, genetischen, psychischen, wirtschaftlichen, kulturellen oder sozialen Identität dieser natürlichen Person sind, identifiziert werden kann (Artikel 4 Ziffer 1 DSGVO).

Der App-Client verzichtet auf eine Registrierung durch die Betroffenen. Mithin ist eine Identifizierbarkeit zum Beispiel durch im Rahmen der Registrierung erhobenen E-Mail-Adressen, Namen, Wohnorte oder sonstiger Daten ausgeschlossen.

Apps können indessen Kennziffern – sogenannte dynamische oder statische Gerätekennungen – auslesen, die es unter Umständen ermöglichen, einen Rückschluss auf die Person zu ziehen. Ein solcher Rückschluss könnte sich ergeben, wenn und soweit derartige Kennungen mit direkten Identifikationsmerkmalen (Name, Adresse, etc.) verknüpft würden oder sich über den App-Client und etwaiger Drittquellen hinweg ein hinreichendes Profil der Person erstellen lässt, sodass eine Identifikation möglich erscheint. Statische Kennziffern, sogenannte statische Gerätekennungen, müssen für den Betrieb des App-Clients indessen nicht erhoben werden. Mithin ist eine Identifikation über derartige statische Kennungen ausgeschlossen. Dynamische Gerätekennungen – wie zum Beispiel Werbe-IDs – werden ebenfalls nicht benötigt, sodass hierüber – auch kein temporäres – hinreichend detailliertes Profil erstellt

werden kann. Es bleibt allenfalls die IP-Adresse als Identifikator. Inwieweit es ausreicht, dass die verantwortliche Stelle rein theoretisch und unter sehr engen rechtlichen Voraussetzungen eine Auflösung der IP-Adresse zu einem konkreten Anschluss verlangen und mithin über diese Adressen gegebenenfalls eine Person identifizieren kann, ist umstritten.[99] Für das Forschungsprojekt wurde ein Streitentscheid vermieden, soweit die datenschutzrechtliche Tragweite durch technische Maßnahmen vermieden werden konnte.

Fraglich ist, ob sonstige Informationen in Summe zu einer hinreichenden Profilbildung genutzt werden können, sodass eine Identifizierbarkeit angenommen werden kann. Hierzu könnte zum Beispiel auch die Kombination der jeweils installierten Apps in ihrer jeweiligen Version gehören. Es ist jedenfalls in Anbetracht der Vielzahl von Apps und des individuellen Update-Verhaltens der Nutzer nicht auszuschließen, dass auf dieser Basis eine Art Fingerabdruck entstehen könnte. Dieser wäre dann wiederum geeignet das Endgerät zu identifizieren, welches möglicherweise unter Hinzuziehung weiterer Daten eine Identifikation ermöglichen könnte. Dieses Szenario wurde im Forschungsprojekt zwar grundsätzlich als unwahrscheinlich eingestuft. Dennoch konnte hier durch technische Maßnahmen der mögliche Personenbezug definitiv ausgeschlossen werden.

5.4.3.4 Vermeidung des Personenbezugs

Ein Personenbezug konnte aus Sicht der Forschungskonsortiums durch technische Maßnahmen vermieden werden.

Bezüglich der IP-Adresse ist deren Erhebung technisch notwendig. Es kann allenfalls sichergestellt werden, dass deren Verarbeitung nur für den kurzen Moment der technischen Notwendigkeit erfolgt und im Weiteren für keinen der Verarbeitungsschritte als ID herangezogen wird.

Nähme man selbst für diesen kurzen Moment einen Personenbezug aufgrund der IP-Adresse an, können die Daten aus Sicht der verantwortlichen Stelle unbrauchbar gemacht werden. Mit der gleichen Technik kann auch der mögliche Personenbezug durch einen theoretischen Fingerabdruck installierter Apps verhindert werden. So ist es nicht erforderlich, dass Apps in Form ihrer Bundle-Identifier in ihrer Gesamtheit an den Server übermittelt werden. Vielmehr können diese in eine Vielzahl– zufälliger – Zusammenstellungen gruppiert werden. Diese Zusammenstellungen werden sodann nicht auf Basis der IP-Adresse weiterverarbeitet, sondern auf Basis einer temporären, rollierenden ID.

Im Rahmen des Forschungsprojekts und für den Demonstrator werden alle Datenverarbeitungen, die einen Personenbezug darstellen können, ohnehin nur optional angeboten. Zudem werden im PGuard-Backend alle potentiellen Daten, die einen Bezug herstellen können, gekürzt und anonymisiert. Falls zu einem späteren Zeitpunkt eine Relation notwendig werden sollte, wurde dazu ein Verfahren evaluiert und implementiert, welches eine pseudonymisierte Benutzerverwaltung ermöglicht. Hierzu Näheres unter 5.3.1.

[99] EuGH Urteil vom 19. Oktober 2016 Az. C-582/14 (Breyer gegen die BRD), berichtigt durch Beschluss vom 6. Dezember 2016.

Gleiche Überlegungen sind natürlich für einen etwaigen Rückkanal anzustellen. Näheres hierzu unter 5.4.5.

5.4.4 Auswertung lokaler Meta-Daten installierter Apps

Im Rahmen des Forschungsprojektes galt es nicht nur die technischen Möglichkeiten zu untersuchen, sondern insbesondere auch diese den Betroffenen in einer Art und Weise zugänglich zu machen, dass Betroffene die Prüfergebnisse verstehen und aus einem App-Client einen individuellen Nutzen ziehen.

Hiermit galt es die Integration von Komfortfunktionen zu untersuchen. Je nach Einschätzung der Verantwortlichkeit entsprechend 5.4.1.1 verlangen bereits einfache Funktionen – wie das Sortieren nach Installationsdatum – datenschutzrechtliche Überlegungen.

Technisch ist es möglich Metadaten der installierten Apps auszulesen. Einige diese Metadaten verlangen, dass dem App-Client spezifische Berechtigungen eingeräumt wurden.[100] Andere Metadaten können grundsätzlich ausgelesen werden.

5.4.5 Rückkanal zur App zur Übermittlung von zentral hinterlegten Informationen

Entweder wird der Abruf der Prüfergebnisse bei Eingang bereits mit anderen Daten verwässert oder die Ergebnisse werden selbst – ähnlich wie die Bundle-Identifier – in kleineren zufälligen Zusammenstellungen übermittelt.

Zudem galt es sicherzustellen, dass ein solcher Rückkanal nicht selbst auf Basis eines Identifikators arbeitet, der auf Dauer eine Personenbeziehbarkeit ermöglicht. Näheres zur Umsetzung findet sich unter 5.4.3.4.

5.4.6 Integration etwaiger Drittdienste

Um einzelne Funktionen anzubieten und Forschungsergebnisse zu generieren kann der der Rückgriff auf Drittdienste nicht gänzlich ausgeschlossen werden.

Hierzu zählen Funktionen wie Push-Benachrichtigungen aber auch Analyse-Tools. Letztere wurden insbesondere im Rahmen der Forschung für erforderlich erachtet, um das Usability-Konzept bei Bedarf hinsichtlich seiner Verständlichkeit und Sachdienlichkeit für die Betroffenen zu testen.

Das Forschungsprojekt hat hierbei – soweit möglich – auf selbst gehostete Services zurückgegriffen. So wurde zum Beispiel für die Analyse der Einsatz von matomo[101] präferiert. Zugleich wurde vermieden mehrere Drittservices für die gleichen Zwecke zu verwenden, soweit die Daten der bereits eingebundenen Dienste für die zweckdienliche Auswertung im Rahmen des Forschungsprojekts genügten.

Ein Großteil dieser Überlegungen verblieb jedoch im Konzeptstadium, da eine aktive Integration im Demonstrator nicht notwendig erschien.

[100] Unter Android zum Beispiel „Usage“.
[101] https://matomo.org/; früher piwik https://matomo.org/blog/2018/01/piwik-is-now-matomo/ .

5.4.7 Zusammenfassung

Von besonderer Bedeutung waren die datenschutzrechtlichen Anforderungen an den App-Client. Die Prüfung der installierten Apps kann unter mehreren Gesichtspunkten datenschutzrechtliche Relevanz aufweisen. Ungeachtet dessen, erfordert der App-Client minimale Datenverarbeitungen. Datenverarbeitungen und Datenzugriffe wurden zudem – soweit möglich – als Opt-In ausgestaltet.

5.5 Sonstige rechtliche Anforderungen an die durch einen App-Client bereitgestellten Funktionen

Neben ohnehin zu berücksichtigenden Pflichtangaben, Impressum, Datenschutzerklärung, ggf. Allgemeine Nutzungsbedingungen, waren im Rahmen der Forschung auch andere Rechtsgebiete, insbesondere die des Wettbewerbs- und Äußerungsrechts zu berücksichtigen. Unmittelbaren Einfluss auf die Forschung hatten insbesondere die letztgenannten Rechtsgebiete, sodass sich die weitergehenden Ausführungen auf die wesentlichen Aspekte in diesem Kontext beschränken.

Die Angaben, Hinweise und Empfehlungen durch den im Rahmen des Forschungsprojekt zu entwickelnden App-Client sind geeignet sich erheblich auf die betroffenen Märkte auszuwirken. Hierbei galt es insbesondere sicherzustellen, ein Abmahn- und Schadenersatzrisiko zu minimieren

Es galt sicherzustellen, dass die Angaben zutreffende Informationen bereitstellen, soweit tatsächliche Verhältnisse beschrieben werden. Soweit Wertungen vorzunehmen waren, wurde darauf geachtet, dass diese Wertungen die rechtlichen Rahmenbedingungen zulässiger Meinungsäußerung nicht überschreiten.

5.5.1 Sicherstellung zutreffender Tatsachen

Zu unterscheiden sind diejenigen Informationen, die Prüfergebnisse darstellen und diejenigen Informationen, die auf Grundlage der Prüfergebnisse dargestellt werden – im Wesentlichen die Inhalte der Informationsblöcke.

Die Inhalte der Informationstexte wurden entsprechend der Darstellung unter 4.8.1 über die gesamte Zeit des Forschungsprojekts optimiert und mit neuen Erkenntnissen angereichert. Hierbei galt es sicherzustellen, dass trotz der notwendigen Simplifizierungen und Verkürzungen der Aussagegehalt weiterhin den Tatsachen entspricht.

Schwieriger verhielt es sich mit den Prüfergebnissen. Eine automatisierte Prüfung birgt stets das Risiko einer gewissen Unschärfe. Diese Unschärfe galt es im Rahmen des Forschungsprojektes zu adressieren. Hierbei wurde auf bestehende Best-Practices zurückgegriffen, um die bestehenden Risiken zu minimieren aber zugleich den Mehrwert einer automatisierten Prüfung Betroffenen zugänglich machen zu können.

Im Wesentlichen wurden nachstehende Mechanismen im Rahmen der Entwicklung berücksichtigt und, soweit im Rahmen des Forschungsprojekts bei der Entwicklung des Demonstrators sachdienlich, implementiert:

- Reproduzierbarkeit der Ergebnisse;
- Transparenz über etwaige grundsätzliche Unschärfen der Prüfergebnisse in der Kommunikation – auch im Rahmen von allgemeinen Nutzungsbedingungen
- Transparenz über die wesentlichen Aspekte, auf die sich die Prüfergebnisse beziehen (App-Version Nummern, Stand der Datenschutzerklärung, etc.)
- Meldefunktionen für vermeintlich unzutreffende Angaben mit anschließender manueller Prüfung

5.5.2 Sicherstellung bezüglich etwaiger Wertungen und Empfehlungen

Neben reinen Tatsachen werden auch Wertungen kommuniziert. Diese reduzieren sich zwar auf Grund des neutralen Grundansatzes des Forschungsvorhabens (siehe 3 und 4.8.2.2) auf ein Minimum; dennoch sind Empfehlungen weniger den Tatsachen zuzuordnen. Gleiches gilt in begrenztem Umfang auch für die Darstellung der Vor- und Nachteile sowie der Klassifizierung als „rote Linien“[102].

Um die Risiken in diesem Bereich gering zu halten, wurde weitestgehend auf bestehende Materialien zurückgegriffen. Dies erschien auch im Sinne einer effizienten und zielgerichteten Forschung sachdienlich, soweit die bereitgestellten Informationen in ihrer Ausgestaltung in das Funktions- und Bedienkonzept integriert werden konnten.

Zurückgegriffen wurde insbesondere auf bestehende Leitfäden und Informationsportale, insbesondere

- mobilsicher.de,
- BSI-für-Bürger,
- Stiftung Warentest,
- Leitfäden der Datenschutzaufsichtsbehörden,
- Verbraucherzentralen

5.5.3 Zusammenfassung

Der App-Client musste auch sonstige rechtliche Anforderungen beachten. Die Entwicklung hat diese Vorgaben ebenfalls berücksichtigt und – soweit möglich – auf Informationen und Materialien Dritter zurückgegriffen.

6 Forschung und Ergebnisse im Zusammenhang sonstiger Zugangsoptionen (Website, Browser-Plugin, etc.)

6.1 Einleitung

Im Rahmen des Forschungsprojekts sind neben dem App-Client auch weitere Lösungen entstanden, die im Wesentlichen auf den Techniken, Erfahrungen und Überlegungen hinsichtlich der Nutzerfreundlichkeit des App-Clients aufbauen.

[102] Siehe 4.8.2.2.

Das ist einerseits die Möglichkeit, beliebige Datenschutzerklärungen auf einer Webseite semantisch analysieren zu lassen, um Nutzerinnen und Nutzern sodann eine Auswertung bereitstellen zu können, siehe auch 6.2.

Andererseits wurde auch ein Browser-Plugin entwickelt, welches die Informationen zu einzelnen Apps vor Installation und im Google Play Store integriert, siehe 6.3.

6.2 Datenschutzerklärungs-Analyzer bzw. „Check Your–APPS" - Webseite zur Analyse von beliebigen Datenschutzerklärung Freitexten

6.2.1 Umsetzung und übergeordnete Entwicklungsschritte

Im Rahmen des Forschungsprojektes stellte sich heraus, dass nicht alle Apps Datenschutzerklärungen in einer Art und Weise bereitstellen, sodass diese vollautomatisiert aufgefunden und sodann analysiert werden können. Zudem ist es auch denkbar, dass Nutzerinnen und Nutzer Apps nicht aus dem Google Play Store beziehen und dennoch Unterstützung bei der Analyse wünschen.

Um diesen Bedarf zu adressieren wurde im Rahmen des Forschungsprojekts eine Lösung entwickelt, die die semantische Analyse auch im Wege einer Webseite zugänglich macht.

6.2.2 Besonderheiten hinsichtlich der Zugänglichkeit, Verständlichkeit und Bedienbarkeit des Datenschutzerklärungs-Analyzers

Bei der Entwicklung der Nutzeroberfläche und des Bedienkonzepts des Datenschutzerklärung-Analyzers wurden insgesamt drei Schritte absolviert:

1. Schritt: Mission und Vision des Datenschutzerklärung-Analyzers und Usability-Leitplanken

Ziel des Datenschutzerklärung-Analyzers ist es, den Nutzerinnen und Nutzern zu ermöglichen, eine beliebige Datenschutzerklärung automatisch auswerten zu lassen. So können Nutzerinnen und Nutzer eine Datenschutzerklärung einer App oder von einer Webseite einfügen und erhalten nach Prüfung durch die DATENSCHUTZscanner-Technologie Ergebnisse zu den Datenverarbeitungen.

Wie bereits in 4.8.1.2.5 und 5.2.2 erläutert, wurde für die DATENSCHUTZscanner-Anwendungen eine Mission und Vision erstellt, die die Zielvorgaben der Anwendungen beschreibt. Sie lautet:

Der DATENSCHUTZscanner bietet Nutzerinnen und Nutzern mehr Transparenz und Kontrolle beim Thema Datenschutz in ihren Apps.

Der Datenschutzerklärungs-Analyzer fügt sich in die Mission und Vision ein, indem er Nutzerinnen und Nutzern ermöglicht, sich transparent über Datenverarbeitungen zu informieren. Zusätzlich bietet er Nutzerinnen und Nutzern Kontrolle, da sie sich vorab mit dem Datenschutzerklärungs-Analyzer über Datenverarbeitungen informieren können und so die Möglichkeit erhalten, sich vor Installation oder Nutzung eines Service zu entscheiden, ob dieser mit ihren Vorstellungen übereinstimmt. Die in der ersten Nutzerbefragung im Jahr 2016 (siehe 4.5.1, Abbildung 9) gewünschte Funktion „vor Installation neuer Apps

über deren Datenverarbeitung informieren“, die innerhalb der neun abgefragten möglichen Funktionen auf dem zweiten Platz bezüglich ihrer Wichtigkeit rangierte, kann mit Hilfe des Datenschutzerklärungs-Analyzers umgesetzt werden. Insofern ist der Datenschutzerklärungs-Analyzer komplementär zur Labormuster-App (vgl. 5.2.5) zu verstehen und erweitert den Funktionsumfang. Insbesondere ist es Nutzerinnen und Nutzern hierüber möglich auch Datenschutzerklärungen manuell prüfen zu lassen, die durch die automatisierten Prozesse – wie in 4.10.2, 4.10.3 sowie 4.10.5 dargestellt – nicht aufgefunden und ausgewertet werden konnten.

Wie auch im Falle des Labormusters war es zusätzlich notwendig bei der Konzeptionierung des Bedienkonzepts und der Usability die identifizierten „Goldstandards“ bzw. Eigenschaften zu berücksichtigen (vgl. 5.2.2). Hierzu zählten[103]:

1. Nützlich
2. Erlernbar
3. Einprägsam
4. Effektiv
5. Effizient
6. Begehrenswert
7. Reizvoll

Ein Vorteil des Datenschutzerklärungs-Analyzers ist die Integration der Anwendung in eine Webseite. Im Vergleich zu einer App können auf Webseiten mehr Informationen untergebracht werden, da die Bildschirmgröße im Vergleich zu einem mobilen Endgerät größer ist. Hierdurch können Texte und Verlinkungen visuell besser umgesetzt werden und das Unterbringen von Detailinformationen, wie der Informationstexte, ist einfacher.

2. Schritt: Konzeptionierung des Bedienkonzepts des Datenschutzerklärungs-Analyzers

Im Rahmen eines Projektworkshops wurde das Bedienkonzept nebst relevanten Funktionen festgehalten. Neben der Kernfunktion, den Nutzerinnen und Nutzern zu ermöglichen eine beliebige Datenschutzerklärung analysieren zu lassen und dieses Ziel in wenigen, intuitiven Schritten zu erreichen, wurden in das Bedienkonzept die Informationstexte integriert. Nach Auswertung durch die DATENSCHUTZscanner-Technologie sollen die Ergebnisse im gewohnten Format, das heißt inklusive der Vor- und Nachteile einer Datenverarbeitung sowie relevanter Handlungsempfehlungen, angezeigt werden.

3. Schritt: Labormuster des Datenschutzerklärungs-Analyzers: Überarbeitung und Umsetzung des Bedienkonzepts

In einem weiteren Projektworkshop wurde die technische Umsetzung des Bedienkonzepts optimiert.

[103] Vgl. Krug, S. (2014). Don’t make me think! Revisited. Web & Mobile Usability; S. 9.

a) Organisation der Datenverarbeitungen nach Clustern

Auch in den Datenschutzerklärungs-Analyzer wurde das Konzept der Informationstexte-Cluster (vgl. Kapitel 4.8.2.3) integriert. So wird bei der Anzeige der Auswertungsergebnisse angegeben, in welchen Clustern eine Datenverarbeitung festgestellt wurde. Durch eine visuelle Abgrenzung der Cluster wird es den Nutzerinnen und Nutzern ermöglicht, sich zielgerichtet über für sie besonders relevante Datenverarbeitungen zu informieren.

b) Hilfetexte zur Bedienung

Um Nutzerinnen und Nutzern die Bedienung des Datenschutzerklärungs-Analyzers zu erleichtern, wurden Begleittexte verfasst, die die Bedienung der Anwendung unterstützen sollen.

c) Nutzungsbedingungen und Transparenz bei Datenschutzpraktiken des Datenschutzerklärungs-Analyzers

Wie auch im Falle des Labormusters soll Transparenz bei der entwickelten Anwendung über die eigenen Datenschutzpraktiken gewährleistet werden. Deshalb wurde eine Datenschutzerklärung verfasst, die es Nutzerinnen und Nutzern ermöglicht, zu verstehen, welche Daten mit dem Datenschutzerklärungs-Analyzer verarbeitet werden. Entsprechend den Ausführungen unter 5.5 zur Adressierung rechtlicher Risiken im Rahmen des Forschungsprojektes für einen App-Client, wurden auch für diesen Dienst Nutzungsbedingungen verfasst.

Abbildung 80 bis Abbildung 82 zeigen den Datenschutzerklärungs-Analyzer. Abbildung 80 stellt die Startseite dar, in deren unterem Bereich sich ein Textfeld befindet, in welchem eine beliebige Datenschutzerklärung eingefügt werden kann. Durch Betätigen des Buttons „Analysieren“ wird die Auswertung durch die DATENSCHUTZscanner-Technologie gestartet und hiernach das Ergebnis, wie in Abbildung 81 dargestellt, angezeigt. Die festgestellten Datenverarbeitungen werden anhand der Überschrift der entsprechenden Informationstexte angezeigt. Diese sind nach Informationstexte-Cluster farblich unterschiedlich dargestellt und mit Bildsymbolen versehen. Durch Anklicken einer festgestellten Datenverarbeitung klappen zusätzliche Informationen der Informationstexte auf. Dies ist in Abbildung 82 dargestellt. Hierbei werden Detailinformationen angeboten und die Konsequenzen der Datenverarbeitung in Form von Vor- und Nachteilen dargestellt. Außerdem werden Handlungsempfehlungen angezeigt. Zusätzlich werden neben den Informationstexten auch die relevanten Textstellen in der Datenschutzerklärung farblich gehighlightet, sodass Nutzerinnen und Nutzer sich darüber informieren können, welcher Aspekt in der Datenschutzerklärung zur Anzeige des Informationstextes geführt hat.

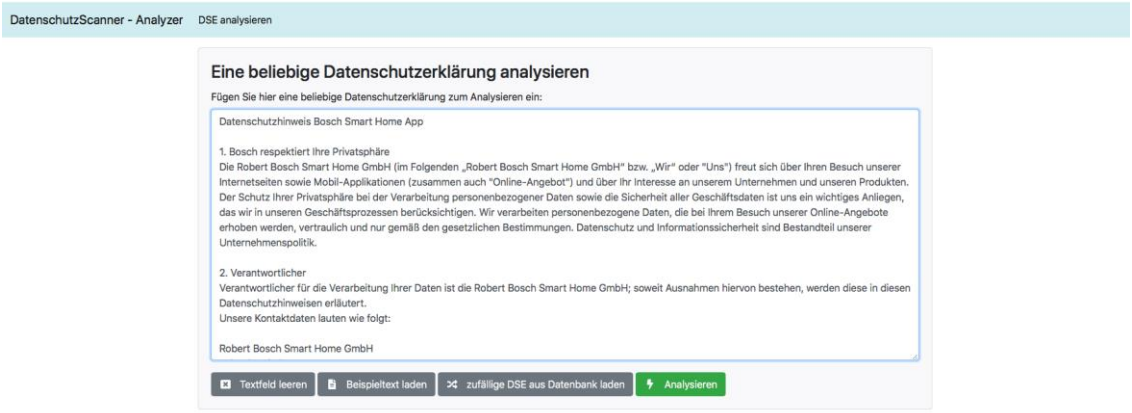


Abbildung 80: Datenschutzerklärungs-Analyzer – Startseite zum Einfügen einer beliebigen Datenschutzerklärung[104]

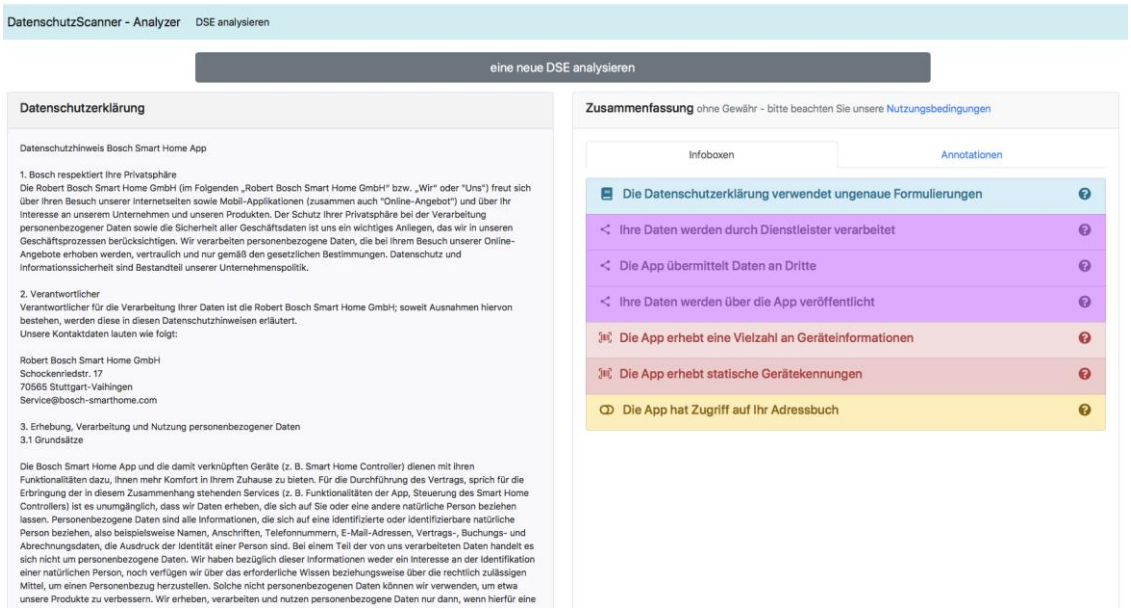


Abbildung 81: Datenschutzerklärungs-Analyzer – Darstellung der Auswertungsergebnisse (Übersicht)[105]

[104] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.
[105] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.

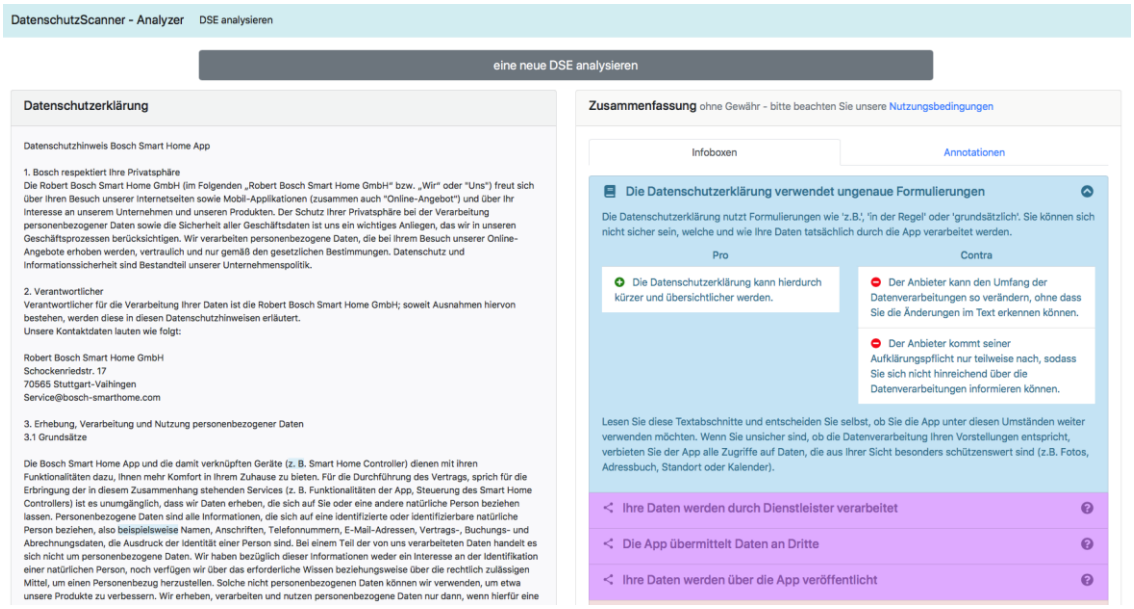


Abbildung 82: Datenschutzerklärungs-Analyzer – Darstellung der Auswertungsergebnisse inklusive Informationstexte[106]

Zum Bedienkonzept des Datenschutzerklärungs-Analyzers liegen keine Ergebnisse aus Nutzertests vor. Die Wirksamkeit der Anwendung ist deshalb nicht bestätigt. Jedoch basiert die Anwendung auf den Ergebnissen der ersten Befragung (vgl. insbesondere 4.5.1 zu gewünschten Funktionen). Außerdem wurde der Datenschutzerklärungs-Analyzer bereits für ein weiteres BMBF-gefördertes Forschungsprojekt eingesetzt und unterstützte in diesem Fall die Auswertung[107].

6.2.3 Technische Umsetzung

Für die "Check Your APPS" Webseite wurde ein eigener Webservice mit Hilfe von Ruby, HTML, CSS (Bootstrap) und JavaScript programmiert.

Auf der Startseite befindet sich ein Textfeld, in welches der Nutzer eine beliebige Datenschutzerklärung kopieren kann. Alternativ kann der Nutzer sich auch einen Beispieltext oder eine andere zufällige Datenschutzerklärung vorschlagen lassen. Mit einem Klick auf „Analysieren" wird der Text an das semantische Backend gesendet und dort analysiert. Anschließend werden die Ergebnisse grafisch aufbereitet.

Die vom Nutzer eingefügte Datenschutzerklärung ist nun auf der linken Seite zu sehen. Auf der rechten Seite kann der Nutzer zwischen der Informationstexte- oder der Annotations-Ansicht wählen. In beiden Fällen werden die gefundenen Textstellen in der Datenschutzerklärung auf der linken Seite farblich hervorgehoben, wenn mit der Maus über einen Informationstext bzw. eine Annotation gefahren wird.

Abbildung 80 stellt die Startseite von „Check your APPS" dar. In dem Textfeld befindet sich eine kurze Datenschutzerklärung, um einen leichten Zugang zu der Technologie zu gewährleisten. Dieser Text kann

[106] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.
[107] Der Einsatz des Datenschutzerklärung-Analyzers fand im Rahmen des Gutachtens „Big Data im Bereich Heim und Freizeit – Smart Living: Status Quo und Entwicklungstendenzen" Kettner et al., 2018 statt, welches für das BMBF-geförderte ABIDA-Projekt (Link: www.abida.de) erstellt wurde.

durch einen Klick auf „Textfeld leeren“ gelöscht und mit einem Klick auf „Beispieltext laden“ wiederhergestellt werden. Mit einem Klick auf „zufällige Datenschutzerklärung aus Datenbank laden“ wird eine zufällige Datenschutzerklärung einer App aus der Datenbank geladen. Diese Texte wurden händisch überprüft, ob es sich wirklich um Datenschutzerklärung von Apps handelt. Der User kann auch einen beliebigen Text in das Textfeld kopieren. Mit einem Klick auf „Analysieren“ wird sodann die semantische Analyse gestartet.

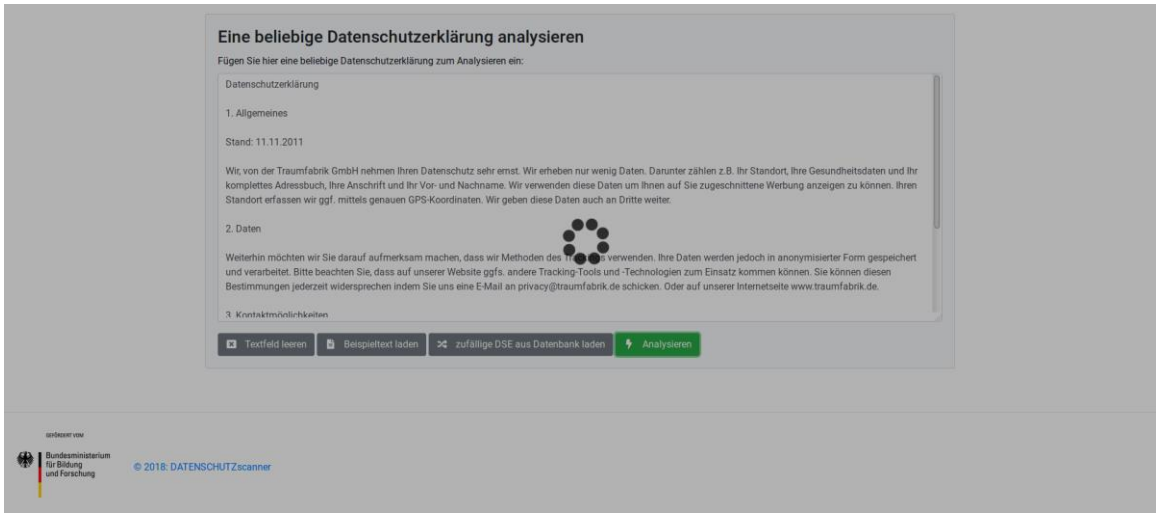


Abbildung 83: Ladebildschirm während der semantischen Analyse

Da die Analyse zwischen 10 und 25 Sekunden dauern kann wird ein Ladebildschirm (Abbildung 83) angezeigt, um einen laufenden Prozess zu symbolisieren. Die Analyse erfolgt direkt im semantischen-Backend.

Nutzerinnen und Nutzern werden sodann die Ergebnisse in Form der Informationstexte[108] als Übersicht angezeigt (Abbildung 81). Auf der linken Seite ist noch einmal die analysierte Datenschutzerklärung sichtbar. Auf der rechten Seite werden nun alle gefundenen Informationstexte angezeigt.

Nutzerinnen und Nutzer können sich auch weitere Details anzeigen lassen (Abbildung 82). Erfolgt ein Klick auf einen Informationstext so werden weitere Informationen zu diesem angezeigt. Weiterhin werden in der linksstehenden Datenschutzerklärung die Textstellen markiert, die zur Anzeige des Informationstexts geführt haben.

[108] Siehe 4.8.2.

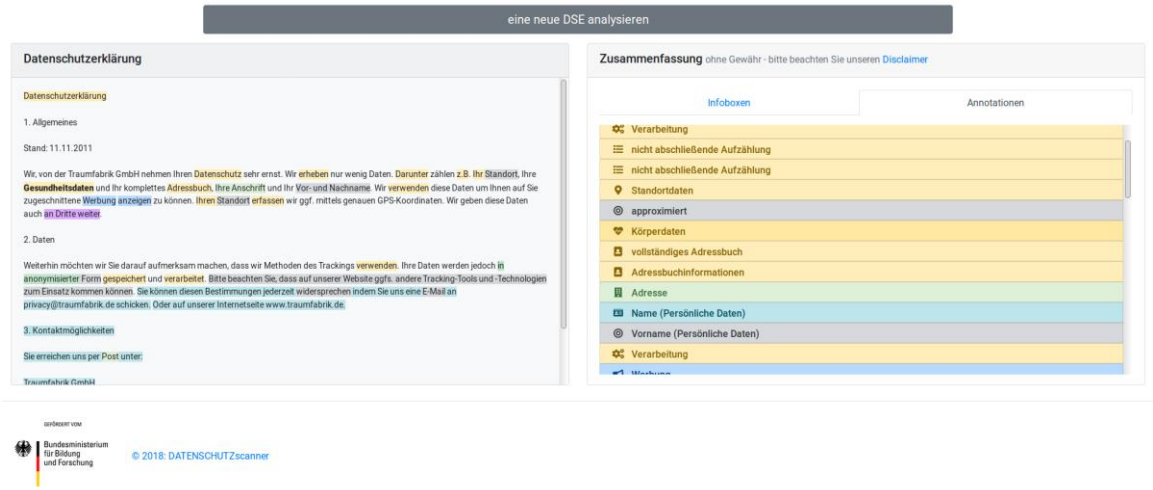


Abbildung 84: Darstellung der Ergebnisse (Annotationsbasis).

Es ist auch möglich, sich alle gefundenen Annotationen, unabhängig von den Informationstexten, anzeigen zu lassen (Abbildung 84). Dazu wird im Reiter die Karte „Annotationen“ ausgewählt. Es werden nun alle gefundenen Annotationen farblich in der Datenschutzerklärung hervorgehoben. Somit werden wichtige Textstellen schnell sichtbar. Fährt man mit der Maus (hover) über eine bestimmte Annotation, so wird diese nochmals fett in der Datenschutzerklärung markiert.

6.3 Browser Plugin zur Anreicherung des Google Play Store um Analyseergebnisse aus dem PGuard Projekt

6.3.1 Umsetzung und übergeordnete Entwicklungsschritte

Im Rahmen der Entwicklung des App-Clients und den damit einhergehenden Befragungen der Nutzerinnen und Nutzer Clients konnte im Forschungsprojekt der Bedarf festgestellt werden, Nutzerinnen und Nutzer bereits vor Installation einer App Analyseergebnisse bereitzustellen, siehe auch 4.5.1.

Die Modifikation der Anzeige in den Appstores auf mobilen Endgeräten ist unzulässig. Eine Bereitstellung der Informationen hätte insofern das Angebot eines eigenen Appstores für die Nutzerinnen und Nutzer bedurft. Dieser Ansatz wurde im Projekt verworfen, Näheres unter 5.4.1.2.

Als Alternative wurde die Entwicklung eines Browser-Plugins angesehen. Hierbei wird die Anzeige des Google Play Stores im Browser modifiziert und um die Analyseergebnisse angereichert.

6.3.2 Besonderheiten hinsichtlich der Zugänglichkeit, Verständlichkeit und Bedienbarkeit eines Browser-Plugins

Bei der Entwicklung der Nutzeroberfläche und des Bedienkonzepts des PGuard Browser-Plugins wurden insgesamt drei Schritte absolviert:

1. Schritt: Mission und Vision des PGuard Browser-Plugins und Usability-Leitplanken

Ziel des PGuard Browser-Plugins ist es, Nutzerinnen und Nutzer kurz und kompakt darüber zu informieren, welche Datenverarbeitungen auf bestimmte, im Google Play Store angebotene Apps, zutreffen. So werden die Informationen der DATENSCHUTZscanner-Technologie durch das PGuard Browser-Plugin in den Google Play Store integriert und angezeigt.

Auch das PGuard Browser-Plugin fügt sich in die in 4.8.1.2.5 beschriebene Mission und Vision der DATENSCHUTZscanner-Anwendungen ein:

Der DATENSCHUTZscanner bietet Nutzerinnen und Nutzern mehr Transparenz und Kontrolle beim Thema Datenschutz in ihren Apps.

Durch Anzeige der relevanten Datenverarbeitungen einer bestimmten App wird Transparenz geschaffen, die durch die Darstellung der gewohnten Informationstexte verstärkt wird. Darüber hinaus wird auch das Ziel der Kontrolle durch Nutzerinnen und Nutzer gefördert. Wie auch beim Datenschutzerklärungs-Analyzer ermöglicht das PGuard Browser-Plugin es sich vorab, das heißt vor Installation oder Nutzung einer App, zu informieren und so unerwünschte Datenverarbeitungen zu vermeiden. So versteht sich das PGuard Browser-Plugin ebenfalls als Komplement zum App-Labormuster.

Die Konzeptionierung der Usability folgt ebenfalls den Usability-Leitplanken, die in 5.2.2 und 6.2.2 beschrieben werden und das PGuard Browser-Plugin soll somit nützlich, erlernbar, einprägsam, effektiv, effizient, begehrenswert und reizvoll sein[109].

Im Unterschied zum Datenschutzerklärungs-Analyzer, der auf einer neuen Webseite erstellt wurde, existieren bezüglich des PGuard Browser-Plugins andere Rahmenbedingungen für die Umsetzung. So wird das PGuard Browser-Plugin in ein vorhandenes Umfeld, den Google Play Store, integriert und es ist deshalb notwendig, dass Nutzerinnen und Nutzer sich in dieser gewohnten Umgebung gut mit Hilfe des PGuard Browser-Plugins orientieren können. Somit ist auch die verfügbare Fläche zur Anzeige der DATENSCHUTZscanner-Ergebnisse begrenzt und die Herausforderung besteht darin, prägnante Informationen intuitiv sichtbar zu machen.

2. Schritt: Konzeptionierung des Bedienkonzepts des PGuard Browser-Plugins

Die Konzeptionierung des Bedienkonzepts fand im Rahmen eines internen Projektworkshops statt. Für die Kernfunktion, kurz und kompakt die relevanten Datenverarbeitungen einer spezifischen App im Google Play Store-Umfeld anzuzeigen, wurde entschieden, die Informationen in die Appliste des Google Play Stores zu integrieren. Die Wahl fiel dabei auf die Anzeige der Anzahl der identifizierten Datenverarbeitungen. Zusätzlich wurde festgelegt, dass die Informationstexte auswählbar sein sollten und die gewohnte Anzeige der Vor- und Nachteile einer Datenverarbeitung sowie Handlungsempfehlungen berücksichtigt werden sollten.

[109] Vgl. Krug, S. (2014). Don't make me think! Revisited. Web & Mobile Usability; S. 9

3. Schritt: Labormuster des PGuard Browser-Plugins: Überarbeitung und Umsetzung des Bedienkonzepts

In einem weiteren Projektworkshop wurde das umgesetzte Bedienkonzept analysiert und weiterentwickelt. Als zusätzliche Funktion im PGuard Browser-Plugin wurde deshalb Folgendes vorgesehen.

a) Sichtbarmachung der „roten Linien“

In der ersten Nutzerbefragung im Jahr 2016 wurden Teilnehmerinnen und Teilnehmer gebeten, mögliche Funktionen einer etwaigen DATENSCHUTZscanner-Anwendung nach ihrer Wichtigkeit zu bewerten (vgl. 4.5.1). Die Befragung ergab, dass die Funktion „besonders kritische Aspekte automatisch anzeigen“ bei Verbraucherinnen und Verbrauchern einen hohen Stellenwert hat, da diese Funktion bezüglich ihrer Wichtigkeit auf dem ersten der neun Ränge lag. Somit wurde festgelegt, dass im PGuard Browser-Plugin eine visuelle Kenntlichmachung von „roten Linien“[110], das heißt besonders kritischer Datenverarbeitungen wie Gesetzesverstöße, umgesetzt werden soll.

Zur Wirksamkeit des PGuard Browser-Plugins liegen bisher keine Ergebnisse aus Nutzertests vor.

6.3.3 Technische Umsetzung

Es wurde jeweils eine Browserextension (PGuard Browser-Plugin) für Chrome und für Firefox programmiert welche funktionsäquivalent sind. Dies deckt einen Browseranteil[111] von insgesamt 64% (= 49% Chrome + 15% Firefox) ab. Der zweitgrößte Browseranteil über 18% von Safari kann größtenteils ignoriert werden, da Safari-User mit einer großen Wahrscheinlichkeit kein Android-Handy nutzen.

Browser Extensions sind aus verschiedenen Komponenten aufgebaut. wie zum Beispiel Hintergrundskripte, Inhaltsskripte, Optionsseiten, GUI-Elemente und verschiedene Logik-Dateien. Diese Komponenten werden mit Hilfe von Webtechnologien wie HTML, CSS und JavaScript nach den jeweiligen Richtlinien von Google und Mozilla programmiert

Das PGuard Browser-Plugin platziert ein kleines Icon neben die Suchleiste im Browser. Über dieses Icon kann die Extension an und abgeschaltet werden. Ist die Extension aktiviert und besucht der Nutzer den Google Play Store, so wird der Google Play Store automatisch mit datenschutzrechtlichen Informationen zu den gerade angezeigten Apps angereichert. Dabei wird, nachdem die Webseite des Google Play Store fertig geladen ist, der HTML-Quelltext des Google Play Stores verändert indem neue HTML-Elemente auf der Webseite eingefügt werden und indem bestehende CSS-Regeln angepasst werden.

Die Extension kann auf verschiedene Events des Browsers reagieren. So filtert sie zum Beispiel beim „onload“ Event, also wenn die Seite fertig geladen ist, ob die aktuell aufgerufene Seite eine Seite des Google Play Stores ist. Falls ja, unterscheidet die Extension, ob die aktuelle Seite eine „Detail-Ansicht“ einer App ist. Ist dies der Fall, so werden die entsprechenden Informationstexte unter der App-Beschreibung eingeblendet. Dabei entsprechen die Informationen exakt den gleichen Texten mit dem gleichen

[110] Siehe 4.8.2.2.
[111] https://www.browser-statistik.de/

Aufbau (First und Second Level, Pro und Contra, Handlungsempfehlungen) wie in der DATENSCHUTZ-scanner App.

Weiterhin wird die aktuelle Seite auf Kachel-Ansichten von Apps hin untersucht. Wurde eine App-Kachel gefunden, so wird über dem Icon ein Balken mit stark zusammengefassten datenschutzrechtlichen Informationen eingeblendet, nämlich die Anzahl der anzuzeigenden Informationstexte. Weiterhin signalisiert die Hintergrundfarbe des Balkens ob unter diesen Funden eine „rote Linie“, siehe hierzu auch 4.8.2.2, ist. Falls ja, so ist der Hintergrund rot. Falls nein, ist er blau. Falls es noch keine Analyseergebnisse zu dieser App gibt, so ist der Hintergrund grau und ein kurzer Hinweistext auf das Nichtvorhandensein von Ergebnissen steht in dem Balken.

Die Informationen zu den Apps werden zurzeit nur aus dem semantischen Backend bezogen. Es stehen also nur Ergebnisse aus der semantischen Analyse, ohne Ergebnisse aus der technischen Analyse, zur Verfügung. Die Extension fragt nur Datensätze zu den Apps an, die gerade auf der aktuellen Seite sichtbar sind. Gibt es zu einer App noch keine Analysedaten so triggert die Extension eine Analyse im Backend.

Weiterhin speichert die Extension die Ergebnisse im LocalStorage des Browsers zwischen um Anfragen an das Backend, und somit Antwortzeiten und Bandbreite, zu vermeiden.

Die Extension verlangsamt den Seitenaufbau nicht, da die Ergebnisse asynchron nachgeladen werden. Der zusätzliche Zeitaufwand, um die Ergebnisse zu laden und anzuzeigen, beträgt ca. 500ms.

Über den App-Icons werden die stark zusammengefassten Ergebnisse der semantischen Analyse angezeigt (Abbildung 85). Die Zahl gibt die Anzahl der gefundenen Informationstexte wieder. Ist der Balken rot, so ist unter diesen Informationstexten mindestens eine rote Linie. Ist der Balken blau so wurde keine rote Linie gefunden. Ist der Balken grau so sind noch keine Analyseergebnisse vorhanden.

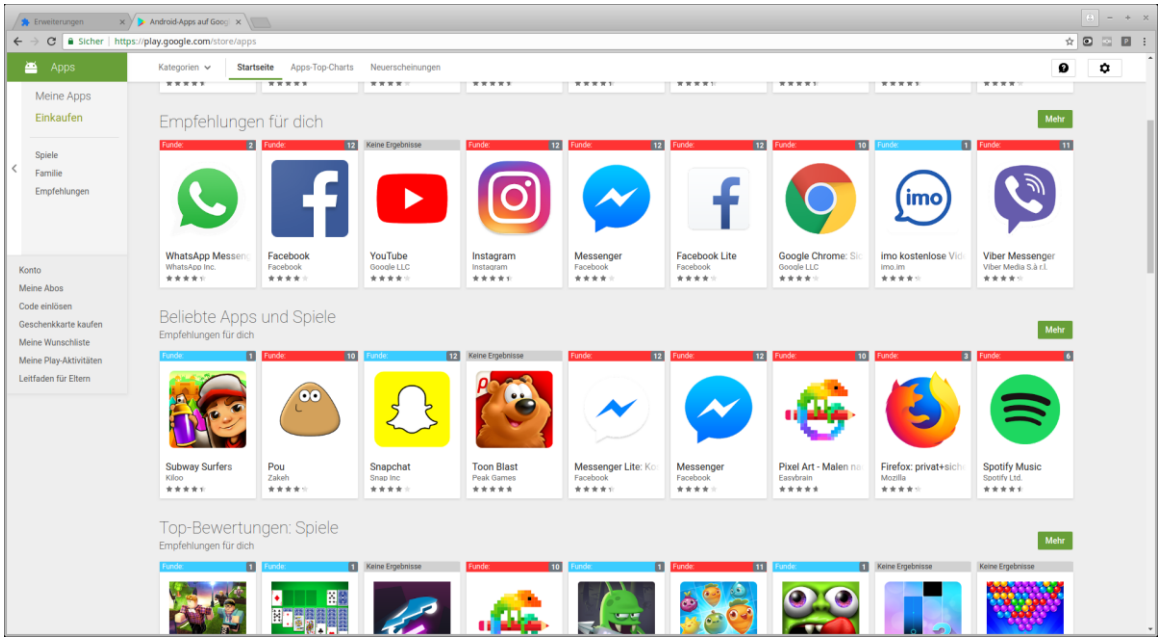


Abbildung 85: Kachel-Ansicht von Apps, hier auf der Startseite des Google Play Stores.[112]

Klicken Nutzerinnen und Nutzer auf einen Balken so erscheint ein PopUp mit detaillierten Informationen zu den Informationstexten; Rote Linien[113] werden farblich hervorgehoben (Abbildung 86).

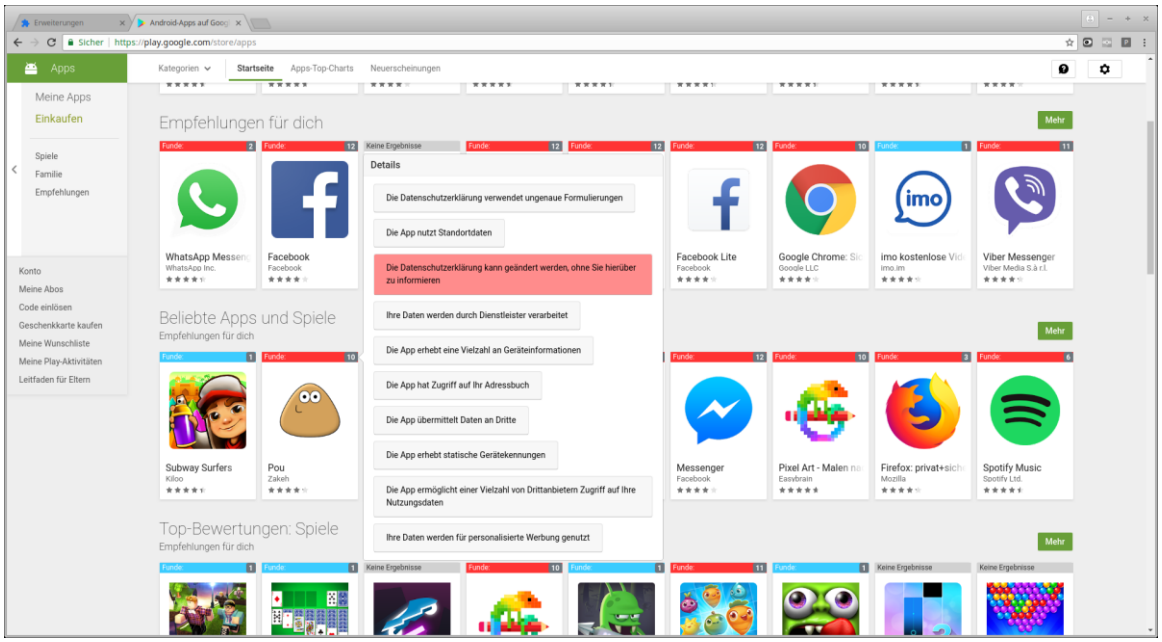


Abbildung 86: Darstellung des PopUps mit weiteren Informationen.[114]

[112] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.
[113] Siehe 4.8.2.2.
[114] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.

Nutzerinnen und Nutzern werden diese Informationen auch in der Detailansicht bereitgestellt. In der Detailansicht werden alle gefundenen Informationstexte unter der App-Beschreibung eingefügt. Rote Linien werden farblich hervorgehoben (Abbildung 87).

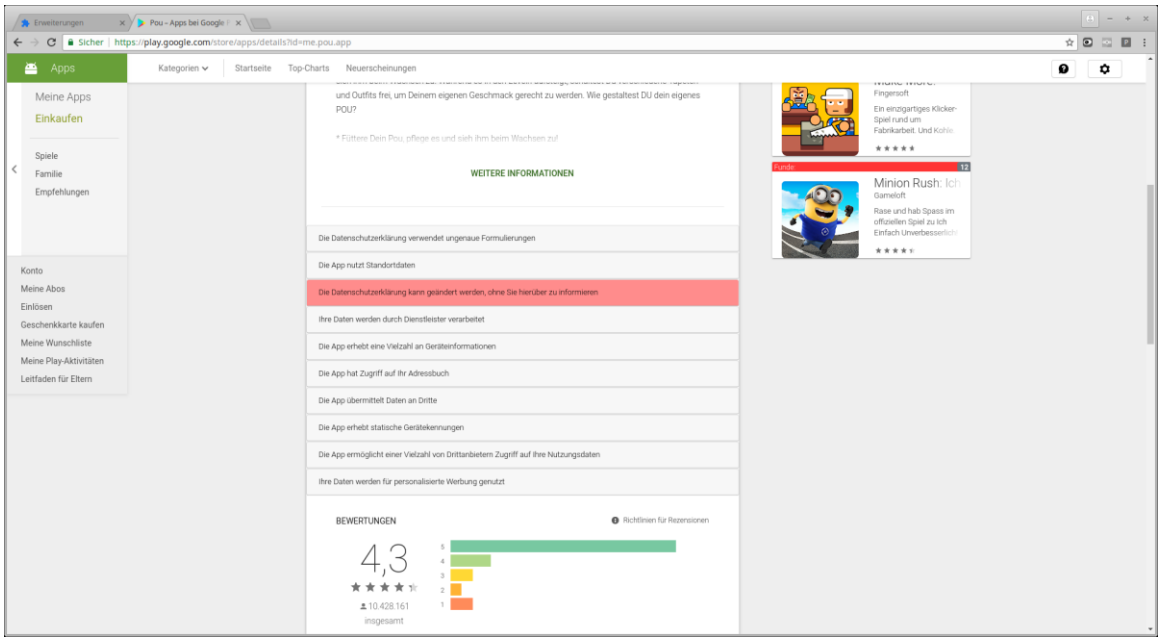


Abbildung 87:Darstellung der Informationen in der Detailansicht.[115]

Weitere Details zu den Informationstexten können durch Nutzerinnen und Nutzer auch hier eingesehen werden (Abbildung 88). Mit einem Klick auf einen Informationstext werden weitere Details zu diesem angezeigt. Die bereitgestellten Informationen sind mit denen, zum Beispiel in der DATENSCHUTZscanner App angezeigten Informationen, identisch.

[115] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.

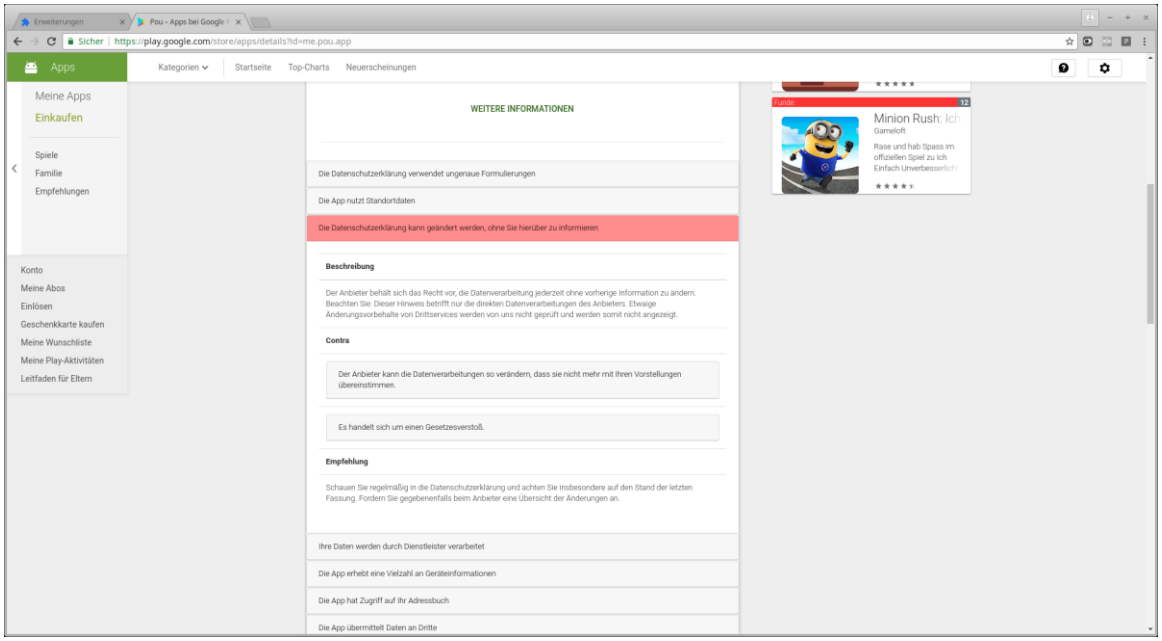


Abbildung 88:Darstellung der Details zu einem Informationstext.[116]

6.4 Datenschutzrechtliche und sonstige rechtliche Anforderungen an durch sonstige Zugangsoptionen bereitgestellten Funktionen

Die sonstigen bereitgestellten Funktionen stellen keine weitergehenden rechtlichen Anforderungen. Zwar sind die gesetzlich zwingenden Informationen unter Umständen an die leicht modifizierte Umgebung anzupassen. Jedoch ergeben sich keine, mit dem konkreten Forschungsvorhaben verbundenen, spezifischen Rechtsfragen.

6.5 Zusammenfassung

Im Laufe des Forschungsprojektes wurden weitere Möglichkeiten erarbeitet, Nutzerinnen und Nutzern die im Forschungsprojekt ermittelten Erkenntnisse zur Verfügung zu stellen. Neben dem App-Client wurde eine Webseite realisiert, mittels welcher Datenschutzerklärungen analysiert werden können. Darüber hinaus wurde ein Browser-Plugin entwickelt, mittels welcher Nutzerinnen und Nutzern im Google Play Store die Prüfergebnisse – also auch vor Installation der jeweiligen App – zur Verfügung gestellt werden können.

[116] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.

7 Verwertungsoptionen

7.1 Marktanalyse

In der ersten Projektphase wurden die Anforderungen für ein Betreiberkonzept ausgearbeitet. Hierbei war es zuerst notwendig zwei Marktakteure zu untersuchen.

1. Wettbewerber, d.h. marktaktive Anbieter von Datenschutzanwendungen
2. Zielgruppe, d.h. die potentiellen Nutzerinnen und Nutzer

7.1.1 Methodik Wettbewerbsanalyse

Ziel der Wettbewerbsanalyse war es, marktaktive Anbieter von Datenschutzanwendungen zu untersuchen. Die methodischen Schritte waren:

1. Identifikation der Wettbewerber und Untersuchung ihrer Geschäftsmodelle.
2. Analyse der angebotenen Funktionen und der Nutzeroberflächen[117] der Wettbewerber.

Die Wettbewerbsanalyse wurde auf Basis einer ausführlichen Matrix durchgeführt, die die wesentlichen Faktoren aufschlüsselt, die für die Nutzung von Selbstdatenschutz-Apps und –Webseiten notwendig sind. Die folgende Tabelle listet die Faktoren der Matrix auf:

Faktor	Beschreibung
Name und Anbieter	• Wer bietet das Produkt an? • Ist der Anbieter bekannt und/oder bietet er auch andere Produkte an? • Wie wurde der Name des Produkts vergeben? Wurden Buzzwords verwendet, die die Funktionalität des Produkts wiederspiegeln?
Beschreibung des Produkts im Appstore bzw. auf der Webseite	• Wie wird das Produkt beschrieben? • Welche Funktionen werden in den Vordergrund gestellt, welche in der Beschreibung nicht erwähnt? • Welche Stichworte werden und welche Tonalität wird verwendet?
Versionen und Preise	• Welche Versionen können genutzt werden? • Gibt es eine Premiumversion? • Wie hoch ist der Preis für die Nutzung? Ergibt sich der Preis aus einem Pauschalbetrag oder einem Abo?
Funktionen	• Welche Funktionen werden (tatsächlich) angeboten? • Werden Nutzerinnen und Nutzer nur bei Öffnen der App bzw. Webseite über die Inhalte informiert oder bedient sich das Produkt an Pop-Ups / visuellen Warnsignalen bei passiver Nutzung?
Präferenzabfrage	• Können die Nutzerinnen und Nutzer des Produkts ihre eigenen Präferenzen einstellen? Wenn ja, wie?
Durchdringung und Nutzerbewertung	• Wie häufig wurde die App bisher heruntergeladen? Wie häufig wurde die Webseite aufgerufen? • Wie sind die Bewertungen in den Appstores? Wie sind die Bewertungen auf den Webseiten (z.B. seiteninterner Blog oder Forum)? • Gibt es wiederkehrende Hinweise / Bemerkungen durch Nutzer?

Tabelle 8: Matrix der Wettbewerbsanalyse

Nach Erstellung der Matrix begann die Recherche in den Appstores bzw. eine Analyse der Webseiten. Die Recherche ergab 13 Wettbewerber, deren Apps (soweit möglich) auf einem Test-Smartphone installiert und getestet wurden. Die Webseiten wurden ebenfalls aufgerufen und analysiert.

[117] Die Untersuchungen der Nutzeroberfläche der wettbewerbsrelevanten Anwendungen sind in die Erarbeitung des Nutzerführungskonzepts eingeflossen. Die Ergebnisse wurden bereits in Kapitel 4 diskutiert und werden deshalb an dieser Stelle nicht erneut aufgeführt.

7.1.2 Methodik Zielgruppen-Analyse

Zweiter Bestandteil der Entwicklung von Anforderungen für ein Betreiberkonzept war die Zielgruppen-Analyse. Diese wurde auf Basis der Befragungsergebnisse der in Kapitel 4 vorgestellten Verbraucherumfrage durchgeführt. Die insgesamt N=1.000 Befragten gaben dabei ihr Interesse an einer Selbstdatenschutz-App oder -Webseite und ihre Zahlungsbereitschaft an. Diese Variablen wurden im Hinblick auf demographische Variablen untersucht und mit dem allgemeinen App-Nutzungsverhalten verglichen.

So flossen auch folgende Forschungsfragen, die sich aus der Literatur- und Wettbewerbsrecherche ergaben, in die Fragebogenerstellung ein und konnten für die Zielgruppen-Analyse verwendet werden:

- Gibt es demografische Effekte (z.B. Alter oder Bildung), die einen Einfluss auf die Nachfrage bzw. die Zahlungsbereitschaft für eine Selbstdatenschutz-App haben?
- Unterscheiden sich die Nachfrage bzw. die Zahlungsbereitschaft zwischen intensiven Smartphone-Nutzern und gelegentlichen Smartphone-Nutzern?

Die Zielgruppen-Analyse wurde mithilfe der Statistiksoftware Stata und SPSS u.a. im Rahmen von Kreuztabellen und multivariaten Analysemethoden durchgeführt.

7.1.3 Ergebnisse der Wettbewerbsanalyse

Insgesamt wurden 13 Angebote analysiert und anhand der Analyse-Matrix ausgewertet. Übergeordnet lässt sich feststellen, dass kein dominanter Anbieter auf dem untersuchten Markt aktiv ist. Die Anwendungen bieten vielfältige Funktionen an, sind jedoch in ihrer Bewertung der datenschutzrechtlichen Aspekte oftmals nicht transparent. Zusätzlich fällt auf, dass viele Anwendungen eine dramatische und gefahrenbetonende Darstellung wählen (Angst- und Sicherheits-Priming[118]).

Abbildung 89 fasst die Eckdaten der untersuchten Anwendungen zusammen. 54% der Anwendungen werden als App im Appstore angeboten und 31% der Anwendungen sind lediglich auf einer Webseite verfügbar. Außerdem betreiben 15% der Anbieter sowohl eine App als auch eine Webseite. Wie Abbildung 90 zeigt, werden lediglich zwei Drittel der Anwendungen für das Betriebssystem Android angeboten. Ein Drittel der Anwendungen bietet den Nutzerinnen und Nutzern ihr Angebot sowohl für Android als auch für iOS an.

[118] Der Begriff Priming bzw. Bahnung bezeichnet in der Psychologie die Beeinflussung der Verarbeitung (Kognition) eines Reizes dadurch, dass ein vorangegangener Reiz implizite Gedächtnisinhalte aktiviert hat.

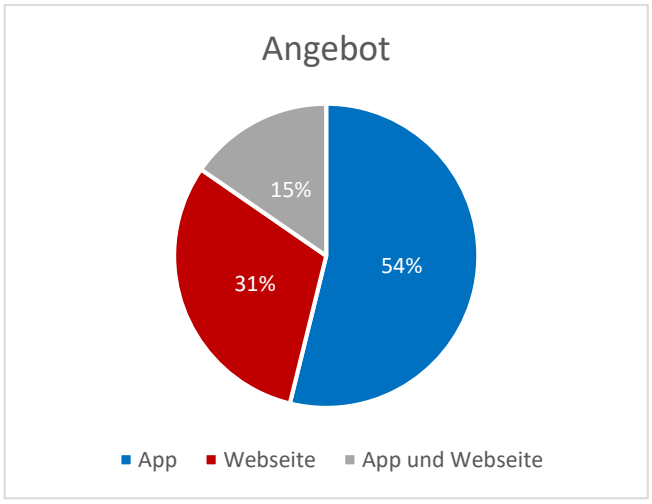


Abbildung 89: Überblick über die untersuchten Anwendungen (N=13)

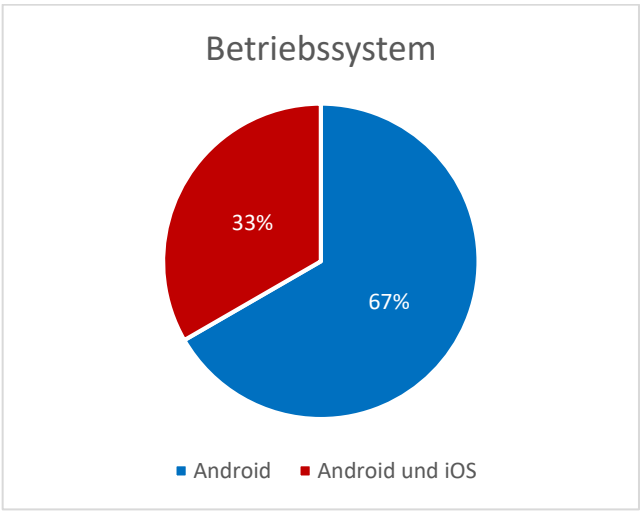


Abbildung 90: Überblick über die untersuchten Anwendungen (N=13)

Die Detailergebnisse, die auf Basis der Analysematrix erstellt wurden, sind in der folgenden Tabelle 9 zusammengefasst:

Faktor	Ergebnis

Name und Anbieter	• Die Namen der Anwendungen decken ein breites Spektrum ab
	  ○ Viele Anwendungen haben englischsprachige Namen, Leitmotiv bei der Namensgebung ist in einzelnen Fällen die Betonung von Sicherheitsbedenken
	• Insgesamt finden sich vielfältige Anbieter auf dem Markt, hierzu zählen
	  ○ Privatwirtschaftliche Unternehmen, Anbieter von etablierter Virensoftware, Prüfinstitutionen und universitätsnahe Einrichtungen
	  ○ Einzelne Anbieter haben einen hohen Bekanntheitswert, da sie sich auf dem Markt für Virensoftware etabliert haben

Beschreibung des Produkts im Appstore bzw. auf der Webseite	• Die Produktbeschreibungen sind teilweise nicht aussagekräftig für die Beschreibung der angebotenen Funktionen
	  ○ Die Funktionen werden selten ausführlich beschrieben; erst bei der tatsächlichen Anwendung werden Funktionsumfang und Bedienkonzepte sichtbar
	• Häufige Verwendung von Angst- und Sicherheits-Priming
	  ○ Die Hälfte der Anwendungen baut auf eine dramatische Darstellung der Datenschutzthematik (Betonung von Gefahren)

Versionen und Preise	• Bei den Versionen ist eine eindeutige Marktfokussierung festzustellen
	  ○ Die Mehrheit der Anwendungen konzentriert sich auf den App-Markt für Android (vgl. Abbildung oben)
	• Die meisten Anwendungen werden kostenlos angeboten
	  ○ Lediglich 3 von 13 Anwendungen bieten eine Premiumversion
	  ○ Die Kosten für die Premiumversion sind entweder monatlich zu tragen (1,45€-2,69€ p.M.) oder als Dauernutzungslizenz verfügbar (3,99€)
	  ○ 2 der 3 Anwendungen mit Premiumversion bieten eine zweiwöchigen Testversion

Funktionen	• Die Kernfunktionen der Anwendungen sind vielfältig – sie decken sich jedoch nicht mit den Funktionen des DATENSCHUTZscanner-Labormusters • Viele Anwendungen setzen auf Berechtigungsmanagement ○ So erlauben die Anwendungen den Nutzerinnen und Nutzern das Management von Berechtigungszugriffen • Einzelne Anwendungen bieten eine DSE-Analyse • Einzelne Anwendungen bieten Malware-Suche als Zusatzfunktion an
Präferenzabfrage	• Präferenzen der Nutzerinnen und Nutzer werden bei keiner Anwendung abgefragt ○ Jedoch ist eine Präferenzeinstellung bei einzelnen Anwendungen möglich, d.h. Bewertungsergebnisse können überschrieben werden

Durchdringung und Nutzerbewertung	• Bei untersuchten Webseiten ist die Nutzungsfrequenz nicht transparent
	• Bei den untersuchten Apps ist die Download-Zahl sehr unterschiedlich
	  ○ Einige Anwendungen haben lediglich 10.000 Downloads, andere über 1.000.000 Downloads
	• Die Nutzerbewertungen der Anwendungen unterscheiden sich stark
	  ○ Zum Teil geben Nutzerinnen und Nutzer positives Feedback und betonen die Nutzungsfreundlichkeit sowie die Wirkung der Anwendungen
	  ○ Bei anderen Anwendungen wird jedoch bemängelt, dass die Anwendungen Angst- und Sicherheits-Priming einsetzen
	  ○ Weiterhin werden die Kosten der Premiumangebote negativ bewertet

Tabelle 9: Ergebnisse der Analysematrix

7.1.4 Ergebnisse der Zielgruppenanalyse

Insgesamt ließ sich aus der Befragung ableiten, dass Verbraucherinnen und Verbraucher ein Interesse an Datenschutz-Anwendungen haben. Informationen zu den Detailergebnissen können in 4.5.3 und Abbildung 10 nachvollzogen werden.

Wie bereits vorgestellt, sind bei der Nachfrage nach Datenschutzanwendungen bestimmte sozio-demographische Attribute relevant. Zwar kann in den Befragungsergebnissen kein Zusammenhang zwischen dem Interesse an eine Datenschutzanwendung und dem Geschlecht festgestellt werden, jedoch existiert ein Alterseffekt. So geben jüngere Teilnehmerinnen und Teilnehmer der Befragung im Vergleich zu älteren häufiger an, ein Interesse an einer Selbstdatenschutz-App zu haben.

Beim Interesse an Datenschutzanwendungen ist außerdem ein weiteres Attribut relevant: So geben Nutzerinnen und Nutzer, die bereits eine oder mehrere Apps auf ihrem Smartphone installiert haben, auch ein größeres Interesse an einer Selbstdatenschutz-App an. Mit der Anzahl der auf dem Smartphone installierten Apps, steigt das Interesse an einer Datenschutzanwendung.

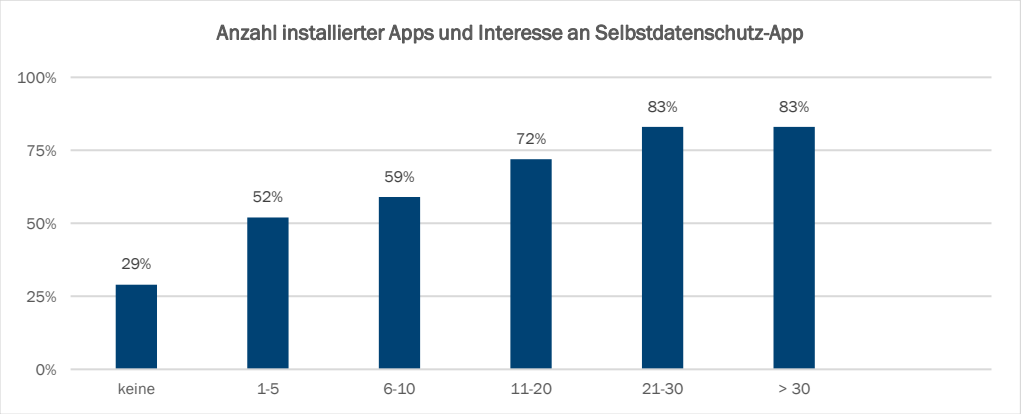


Abbildung 91: Nachfrage nach Datenschutz-App und Anzahl installierte Apps

Wie in 4.5.3 wurde in der Befragung eine positive Zahlungsbereitschaft für eine Datenschutz-App festgestellt. So gaben mehr als zwei Drittel der Befragten eine Zahlungsbereitschaft über 0€ an. Die Zahlungsbereitschaften spannten sich von 0€ bis 400€, mit einem Mittelwert von 5,85€ und einem Median von 3€.

Bei der Zahlungsbereitschaft wurden keine sozio-demographischen Unterschiede festgestellt, so dass die Zahlungsbereitschaft nicht Zielgruppenabhängig zu sein scheint.

7.1.5 Zusammenfassung

Die Ergebnisse zeigen, dass der Markt für Datenschutzanwendungen divers ist. Die Untersuchung der Wettbewerber konnte demonstrieren, dass Anbieter vor allem ein Berechtigungsmanagement für das Smartphone anbieten, nicht jedoch den vollen Funktionsumfang der konzipierten DATENSCHUTZscanner-Anwendungen umfassen. Zusätzlich wurde gezeigt, dass Verbraucherinnen und Verbraucher ein grundsätzliches Interesse an Datenschutzanwendungen haben, dieses jedoch stark von der Zielgruppe abhängt. So sind insbesondere Nutzerinnen und Nutzer, die eine hohe Anzahl an Apps besitzen, an einer Selbstdatenschutzanwendung interessiert. Darüber hinaus kann ein Alterseffekt beobachtet werden. Diese Aspekte sollten somit bei der Bearbeitung des Verwertungskonzepts berücksichtigt werden.

7.2 Untersuchung etwaiger Verwertungsmodelle

Auf Basis der gewonnenen Erkenntnisse der Markt- und Zielgruppenanalyse galt es zu untersuchen, inwieweit die Forschungsergebnisse verwertet werden können. Hierbei ist zu berücksichtigen, dass im Rahmen des Projekts keine finalen Business-Konzepte entwickelt werden konnten und sollten. Außerdem wurden nicht zu allen Verwertungsoptionen derart tiefgreifende Markt- und Zielgruppenanalysen erstellt, wie für das primäre Forschungsergebnis, den DATENSCHUTZscanner App-Client. Grundsätzlich sind einer Verwertung sowohl die konkreten Forschungsergebnisse (DATENSCHUTZscanner[119], Weboberfläche[120], Browser-Plugin[121]) als auch die dahinterliegenden Analyse-Techniken zugänglich.

7.2.1 Grundlegende Annahmen über Entwicklungskosten bis zur Marktreife

Im Rahmen des Forschungsprojektes durfte nur ein sogenanntes Labormuster entwickelt werden. In dieses Labormuster sind alle Elemente eingeflossen, die im Rahmen des Forschungsprojektes eines proof-of-concepts oder einer steten Optimierung (Nutzerführung) bedurften. Ebenso verhielt es sich mit der Entwicklung der Analysetechniken. Die Erkenntnisse wurden in die Analysetechniken eingepflegt, jedoch stets vor dem Forschungshintergrund und der Abwägung zwischen Ressourcenaufwand und Mehrwert für die weitergehende Forschung.

Hinsichtlich der Kosten zur Entwicklung eines marktfähigen Produkts sind daher sowohl im Rahmen der Nutzerführung und Nutzererfahrung als auch im Rahmen der Stabilisierung und Weiterentwicklung der Analysetechniken hinsichtlich einer signifikant höheren Anzahl durchzuführender Analysen je Zeiteinheit Entwicklungsaufwände zu erwarten. Vor einer gut vorbereiteten Produkteinführung stehen auch mehrere Wiederholungsvorgänge von geschlossenen und offenen Nutzertests, deren Erkenntnisse in weitere Produktiterationen eingearbeitet werden. Das hat zum Ziel, ein Produkt zu schaffen, das den Nutzerinnen und Nutzern einen Mehrwert bietet.

Zudem handelt es sich um Analysen dynamischer Inhalte. Beide Analysen müssen stets an geänderte Umsetzungen der zu analysierenden Apps angepasst werden. Es ist zudem damit zu rechnen, dass sich auch bewusste Obfuscationen in der Praxis entwickeln, um ansonsten kritische Ergebnisse derartiger automatisierter Analysen zu vermeiden. Ein dem Forschungskonsortium ähnlich interdisziplinäres Entwicklerteam dürfte daher auch weiterhin erforderlich sein.

Als Kostenfaktoren sind demnach zu erwarten:

- Personalkosten (Juristen, Entwickler – für Serverinfrastruktur, App-Entwicklung und sonstige Frontend-Applikationen, technische Analysen, und semantische Analysen –, UX-Designer, Verhaltensforscher)
- Betriebskosten Infrastruktur (Server und Betriebsräume)
- Marketing und Verkauf

Um die Betriebskosten decken zu können, sind drei Zielmärkte denkbar, in denen Einnahmen zur Deckung der Kosten generiert werden können: Business-to-Consumer (B2C), Business-to-Business (B2B)

[119] Siehe 5.
[120] Siehe 6.2.
[121] Siehe 6.3.

und Business-to-Government/Authority (B2G/A). Eine Verwertung in Form eines White-Labeling oder Lizenzmodells wird ungeachtet des Zielmarktes der Produkte, in die ein etwaiges Forschungsergebnis eingebunden würde, im Rahmen des B2B-Marktes betrachtet. Diese werden im Folgenden näher erläutert.

7.2.1.1 B2C-Verwertung

Für eine Verwertung gegenüber Verbraucherinnen und Verbrauchern eignen sich nur die umfassenden Forschungsergebnisse (wie der DATENSCHUTZscanner App Client), nicht jedoch die dahinterliegenden Analysetechniken als eigenständiges Produkt. Dies entspricht auch dem Ziel des Forschungsvorhabens, Lösungen im Bereich des Selbstdatenschutzes für Verbraucherinnen und Verbraucher zu entwickeln.

Diese Verwertung kann in unterschiedlichen Varianten erfolgen: Von Premium-Modellen über Freemium-Modelle bis hin zu kostenlosen Verwertungen.

7.2.1.1.1 Premium-Modell

Die Verbraucherbefragung ergab, dass Verbraucherinnen und Verbraucher grundsätzlich bereit sind, für einen besseren Selbstdatenschutz zu bezahlen. Bei der Umsetzung sind mehrere Bezahloptionen denkbar:

1. ein Abonnement-Modell, bei dem Verbraucherinnen und Verbraucher beispielsweise monatlich für ihre Nutzung ein Entgelt bezahlen und das vor allem aus dem Bereich der IT-Sicherheitslösungen bereits bekannt ist,
2. ein Kauf-Modell, bei dem Verbraucherinnen und Verbraucher durch eine einmalige Zahlung den Zugang zum Produkt erhalten.

Darüber hinaus sind auch Mischoptionen aus dem Abo- und Kauf-Modell denkbar, beispielsweise in Form eines zunächst zeitlich nahezu unbegrenzten Nutzungsrechts für eine bestimmte Version mit definierten Features inklusive Sicherheitsupdates für einen definierten Zeitraum; neue Features und Überarbeitungen in Form neuer Versionen würden dann wiederum kostenpflichtig bereitgestellt werden.

Das letzte zuvor dargestellte Mischmodell ist das klassische Software-Modell, das Verbraucherinnen und Verbraucher aus Computertagen kennen. Indessen wird auch dort dieses Modell zulasten eines Abonnement-Modells verdrängt. Die Bereitstellung von Software als Software-as-a-Service stellt im Ergebnis sowohl für Anbieter als auch Verbraucherinnen und Verbraucher häufig die für beide Seiten bessere Option dar. So können neue Funktionen, Fehlerbehebungen und Sicherheitspatches leicht zentral eingespielt werden und zudem entfallen die eigenen Kosten für Aufbau und Erhaltung von Knowhow zum Betrieb der Software für die Nutzerinnen und Nutzer.

Eine Abonnement-Lösung würde auch mit der Notwendigkeit korrespondieren, regelmäßigen Einnahmen zu erzielen, um den Mehrwert der Verwertungsobjekte aufrecht erhalten zu können. Fraglich ist indessen, ob die notwendigen Umsätze bei akzeptablen Abonnement-Kosten für Verbraucherinnen und Verbraucher erzielt werden könnten. Hier wird näher zu untersuchen sein, wie sich Abonnement-Preis und Nutzerzahlen zueinander verhalten.

Zu berücksichtigen wird auch sein, dass Verbraucherinnen und Verbraucher zwar eine Zahlungsbereitschaft für die Verwertungsobjekte angegeben haben, diese aber mit anderen Ausgaben in diesem Zusammenhang rivalisiert. Die Kosten für die Verwertungsobjekte müssen sich somit auch in die allgemeine Zahlungsbereitschaft im Bereich Datenschutz und Datensicherheit integrieren lassen.

Vor diesem Hintergrund erscheint ein Modell, allein basierend auf Einmalzahlungen, nur schwer möglich, sodass weitere Untersuchungen nicht angestellt wurden.

7.2.1.1.2 Freemium

Auch denkbar sind etwaige Freemium Modelle. Hierbei würden die Verwertungsobjekte grundsätzlich kostenfrei angeboten. Allerdings würden bestimmte Premium-Funktionen bepreist. Ein solches Modell hat den Vorteil, dass eine hohe Marktdurchdringung, mangels Kostenhürde, leichter zu erreichen ist. Eine sachdienliche Auswahl der Premium-Funktionen könnte zudem dennoch zu signifikanten Umsätzen führen. Ob diese aber allein ein tragfähiges Modell darstellen können, galt es als fraglich einzustufen.

7.2.1.1.3 Free

Gerade im Appmarkt sind „kostenlose“ Angebote üblich. Eine solches „Free“-Modell stellt eine etwaige Verwertung der Verwertungsobjekte allerdings vor die Herausforderung, dass die für die stetige Entwicklung notwendigen Einnahmen auf anderem Wege generiert werden müssten. Hierbei sind folgende Optionen denkbar:

1. Werbefinanzierung
2. Verwertung von Nutzerprofildaten
3. Verwertung aggregierter Daten über das Nutzer- und Datenschutzverhalten
4. Abschreibung als „Werbemaßnahme“ für eine etwaige B2B-Verwertung
5. Spenden und öffentliche Mittel

Im Rahmen des Forschungsprojektes wurden allenfalls die letzten beiden Optionen für sachdienlich erachtet. Die anderen Optionen sind zwar rechtlich und wirtschaftlich nicht verwerflich oder gar verboten, allerdings steht eine derartige Verwertung mit den Zielen der Verwertungsobjekte klar im Widerspruch. Mithin wurde eine solche Verwertung als die Glaubwürdigkeit eines solchen Vorhabens schädlich eingestuft, welches den gesamten Mehrwert einer Verwertung negieren würde.

Die Optionen 4 und 5 erscheinen denkbar. Dabei ist Option 4 selbst beeinflussbar und lediglich davon abhängig, ob entsprechende Verwertungen im B2B-Segment stattfinden werden; Option 5 ist indessen nur geringfügig beeinflussbar und wird daher lediglich als Option aufgeführt ohne tiefergehende Analyse.

7.2.1.2 B2B-Verwertung

Datenschutzrechtliche Prüfungen haben auch im Geschäftskundenumfeld eine steigende Relevanz. Ähnlich der bisherigen IT-Sicherheitstechnischen Compliance wird sich – spätestens unter der Datenschutzgrundverordnung – eine datenschutzrechtliche Compliance-Struktur in Unternehmen entwickeln (müssen). Die im Rahmen des Forschungsprojekts entwickelten Technologien könnten auch in einem solchen Kontext Anwendung finden. Hierbei wären die durch die Analysen ermittelten Informationen in

einer anderen Art und Weise aufzubereiten, als dies bisher für Verbraucherinnen und Verbraucher stattfindet. Es ist jedoch auf lange Sicht von einem hohen Grad synergetischer Effekte auszugehen.

Denkbar erscheinen somit insbesondere nachstehende Modelle:

1. Lizensierung zur Integration in unternehmensinterne (datenschutzrechtliche) Compliance-Strukturen
2. Aufbau und Verwertung der durch die Analysetechnik möglichen Erkenntnisse in ein Beratungsangebot
3. Verwertung der Analysetechnik in Form von Auftragsanalysen
4. Aufbau und Angebot einer eigenen datenschutzspezifischen App-Compliance-Struktur
5. White-Labeling
6. Vollständiger Verkauf der Technologien

Die genannten möglichen Verwertungsoptionen wurden im Rahmen des Forschungsprojektes mit unterschiedlichen Vertretern der Wirtschaft diskutiert und verifiziert. Die jeweilige Verwertung und die konkrete Konzeptionierung sind indessen von einer Vielzahl von Faktoren abhängig.

Im Rahmen der Lizensierung, White-Labeling oder des vollständigen Verkaufs der Technologie erscheint es aus Sicht des Forschungskonsortiums wichtig, dass die geplanten Einsatzszenarien durch die Erwerber mit den Zielen des Forschungsvorhabens kompatibel sind. Mithin sind diese Optionen selbst wiederum davon abhängig, im Unternehmensumfeld einen hinreichenden Mehrwert bei der Integration in etwaige Endprodukte oder Entwicklungsrichtlinien zu generieren. Hierbei sind insbesondere Verwertungsoptionen als Cross-Selling beziehungsweise als Ergänzung zu bestehenden und funktionierenden B2B-Geschäftsmodellen im Bereich MDM – Mobile Device Management – denkbar.

Eine besonders interessante Verwertungsoption könnte sich durch White-Labeling und / oder Lizenzmodelle ergeben, die die Technologie standardmäßig auf den Endgeräten der Verbraucherinnen und Verbraucher bereitstellt; denkbar wären hier Verwertungen zusammen mit Endgeräteherstellern oder Telekommunikationsanbietern, letztlich wohl auch Betriebssystemanbietern. Hierdurch wäre eine signifikante Marktdurchdringung sichergestellt, auf deren Basis weitere Verwertungsmodelle entwickelt werden könnten. Insbesondere würden aber die erforschten Mehrwerte flächendeckend für Verbraucherinnen und Verbraucher zugänglich.

Lizenzmodelle könnten sich auch im Bereich der Anbieter von IT-Sicherheitslösungen ergeben, die über diesen geübten Vertriebskanal die im Forschungsprojekt entwickelten Mehrwerte an Verbraucherinnen und Verbraucher anbieten.

In der betrieblichen Praxis bereits als tragfähig erwiesen haben sich im Bereich der IT-Sicherheit entsprechende Analysen zur Integration in unternehmensinterne Compliance-Strukturen. In Gesprächen mit entsprechenden Vertretern der Wirtschaft konnte ein entsprechender Bedarf auch für datenschutzrechtliche Analysen ermittelt werden. Dies als Grundannahme vorausgesetzt ergibt sodann auch weitere mögliche Verwertungsoptionen, wie zum Beispiel Auftragsanalysen für konkrete Testreihen oder aber auch Entwickler, die eine Compliance bereits im Rahmen der Entwicklung sicherstelle möchten (privacy- und compliance-by-design).

7.2.1.3 B2G/A

Ebenfalls denkbar erscheint eine Verwertung der Technik als auch der konkreten Verwertungsobjekte im B2G/A-Markt, das heißt für Datenschutzaufsichtsbehörden oder Kartellbehörden.

Einerseits besteht die Möglichkeit, die konkreten Verwertungsergebnisse im Rahmen verbraucherpolitischer Bildungsmaßnahmen zu integrieren.

Andererseits bestehen auch Möglichkeiten, die Technik durch entsprechende Ministerien und öffentliche Aufsichtsstellen zu nutzen, um etwaige Rechtsverstöße und Marktentwicklungen besser zu beobachten und – soweit notwendig – korrigierende Maßnahmen zu ergreifen.

Der Einsatz der Technik in öffentlichen Stellen zur Sicherung der Compliance ist ebenfalls denkbar, wird hier aber als eine Verwertung im B2B-Markt betrachtet.

7.2.2 Zusammenfassung

Grundsätzlich lässt sich feststellen, dass es eine Vielzahl von Verwertungsoptionen für die im Forschungsprojekt entstandenen Ergebnisse und deren Potentiale gibt. Kritisch zu betrachten ist die alleinige Tragfähigkeit nur einer einzigen Verwertungsoption, insbesondere, im B2C-Markt. Hierbei stellt vor allem die Notwendigkeit der Finanzierung stetiger Entwicklung die größte Herausforderung dar.

8 Abschluss

Im Rahmen des Forschungsprojekt konnten zielführende Techniken entwickelt werden, die die vom Forschungsprojekt intendierten Ziele erreichen können.

- Im Bereich der Verständlichkeit konnten durch eine standardisierte Aufbereitung und intensive Testung mit Verbraucherinnen und Verbrauchern deutliche Fortschritte im Vergleich zu bisherigen Ansätzen erreicht werden.
- Im Bereich der technischen Analyse konnten Methoden entwickelt werden, die es mit einem hohen Grad der Automatisierung ermöglicht für Verbraucherinnen und Verbraucher relevante Informationen zu den Datenverarbeitungsvorgängen von Apps mit einem hinreichenden Grad der Verlässlichkeit zu erhalten.
- Im Bereich der semantischen Analyse konnte nachgewiesen werden, dass die für Verbraucherinnen und Verbraucher relevanten Informationen automatisiert und mit einem hinreichenden Grad der Verlässlichkeit ermittelt werden können.

Alle im Forschungsprojekt entwickelten Ergebnisse wurden vor dem Hintergrund eines sich dynamisch entwickelnden App-Markts und Rechtsrahmens entwickelt. Die gewählten Analyse-Methoden und im Rahmen der Analysen ermittelten Prüfergebnisse sind sowohl flexibel, um an neue Entwicklungen und weitere relevante Aspekte angepasst zu werden, als auch bereits jetzt so umfangreich, dass bereits jetzt auf Basis der durch diese Methoden ermittelten Ergebnisse hinreichend verlässliche Aussagen auch über weitergehende – zukünftig relevante – Aspekte getroffen werden könnten.

Die Informationstexte folgen einem standardisierten Aufbau, sodass weitere Aspekte leicht abgebildet werden können. Soweit möglich wird zudem auf bestehende Empfehlungshinweise durch Verweise zurückgegriffen, sodass bei geänderten Rahmenbedingungen nicht ständig Anpassungen der entsprechenden Texte im PGuard-Backend erfordern.

Die Prüfergebnisse werden in einem zentralen PGuard-Backend zusammengeführt und vorgehalten. Hierdurch besteht eine Flexibilität, die Ergebnisse in mehrere Aufbereitungsformen – wie zum Beispiel die drei Labormuster – zu überführen. Ebenfalls können hierdurch auch weitere Erkenntnisquellen hinzufügefügt werden, um die Qualität der Ergebnisse weiter zu verbessern.

9 Appendix

9.1 Informationstexte

9.1.1 Grundsätzlicher Aufbau und allgemeine Hinweise

Überschrift	
Detailinformation zur Überschrift	
⊕ Erläuterungen der Konsequenzen 1 ⊕ Erläuterungen der Konsequenzen 2	⊖ Erläuterungen der Konsequenzen 1 ⊖ Erläuterung der Konsequenzen 2
⊕ Detailinformationen zu Konsequenz 1 ⊕ Detailinformationen zu Konsequenz 2	⊖ Detailinformationen zu Konsequenz 1 ⊖ Detailinformationen zu Konsequenz 2

Handlungsoptionen

Rote Linien[122] sind durch einen roten Rahmen gekennzeichnet. Zudem erfolgt die Sortierung anhand der Cluster[123].

[122] Siehe 4.8.2.2.
[123] Siehe 4.8.2.3.

9.1.2 Datenschutzerklärungen

Aspekte, die die Datenschutzerklärungen der entsprechenden Anwendungen betreffen.

Die App stellt keine Datenschutzerklärung bereit[124]	
Anbieter sind gesetzlich verpflichtet, Sie über die Verarbeitung personenbezogener Daten zu informieren. Dies erfordert grundsätzlich eine Datenschutzerklärung.	
	⊖ Es handelt sich um einen Gesetzesverstoß. ⊖ Ihre Daten werden verarbeitet, ohne dass Sie sich hierüber informieren können.
	⊖ Dies hat nicht zwingend einen Missbrauch Ihrer Daten zur Folge. ⊖ Dies hat nicht zwingend einen Missbrauch Ihrer Daten zur Folge.

Verbieten Sie der App alle Zugriffe auf Daten, die aus Ihrer Sicht besonders schützenswert sind (z.B. Fotos, Adressbuch, Standort oder Kalender).

Informieren Sie sich möglichst vor der ersten Nutzung durch Medien, ob die Datenverarbeitung der App mit Ihren Vorstellungen übereinstimmt.

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

Die App benennt keine Kontaktmöglichkeit für datenschutzrechtliche Anliegen[125]	
Sie haben eine Vielzahl von Betroffenenrechten - insbesondere die Rechte auf Auskunft, Berichtigung und Löschung. Der Gesetzgeber verlangt, dass diese Rechte leicht und ohne besondere Hürden ausgeübt werden können. Daher müssen Anbieter eine direkte Kontaktmöglichkeit angeben.	
	⊖ Es handelt sich um einen Gesetzesverstoß. ⊖ Ihre datenschutzrechtlichen Anliegen erreichen möglicherweise nicht die zuständige Abteilung.
	⊖ Dies hat nicht zwingend einen Missbrauch Ihrer Daten zur Folge.

Stellen Sie Ihre datenschutzrechtlichen Anliegen über die allgemeinen Kontaktmöglichkeiten des Anbieters, z.B. diejenigen aus dem Impressum. Weisen Sie bereits im Betreff eindeutig auf den datenschutzrechtlichen Bezug hin.

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

[124] Wurde als Rote Linie im Sinne des 4.8.2.2 definiert.
[125] Wurde als Rote Linie im Sinne des 4.8.2.2 definiert.

Die Datenschutzerklärung kann geändert werden, ohne Sie hierüber zu informieren[126]	
Der Anbieter behält sich das Recht vor, die Datenverarbeitung jederzeit ohne vorherige Information zu ändern. Beachten Sie: Dieser Hinweis betrifft nur die direkten Datenverarbeitungen des Anbieters. etwaige Änderungsvorbehalte von Drittservices werden von uns nicht geprüft und werden somit nicht angezeigt.	
	⊖ Es handelt sich um einen Gesetzesverstoß. ⊖ Der Anbieter kann die Datenverarbeitungen so verändern, dass sie nicht mehr mit Ihren Vorstellungen übereinstimmen.
	⊖ Dies hat nicht zwingend einen Missbrauch Ihrer Daten zur Folge. ⊖ Soweit der Anbieter die Datenverarbeitungen der App reduziert oder die Änderung keine Datenverarbeitungen betrifft, kann dies im Ausnahmefall keinen Gesetzesverstoß darstellen.

Schauen Sie regelmäßig in die Datenschutzerklärung und achten Sie insbesondere auf den Stand der letzten Fassung.

Fordern Sie gegebenenfalls beim Anbieter eine Übersicht der Änderungen an.

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

[126] Wurde als Rote Linie im Sinne des 4.8.2.2 definiert. Es ist allerdings festzuhalten, dass hierbei die Datenverarbeitung betreffende Änderungen – also wesentliche Änderungen – gemeint sind.

Die App verarbeitet Daten, die ausdrücklich ausgeschlossen wurden[127]	
Unsere technische Analyse zeigt, dass die App personenbezogene Daten verarbeitet, die in der Datenschutzerklärung ausdrücklich ausgeschlossen wurden. Hierunter fällt auch eine abweichende Form der konkreten Verarbeitung, z.B. wenn der Anbieter angibt, Daten pseudonymisiert zu erheben, aber trotzdem Ihre Daten im Klartext versendet.	
	⊖ Es handelt sich um eine Täuschung und einen Gesetzesverstoß. ⊖ Der Anbieter ist unseriös. ⊖ Es besteht eine erhöhte Missbrauchsgefahr für Ihre Daten.
	⊖ ⊖ ⊖ Das konkrete Risiko hängt von den betroffenen Daten und dem konkreten Missbrauch ab.

Deinstallieren Sie die App.

Wenn sie das nicht möchten, verbieten Sie der App alle Zugriffe auf Daten, die aus Ihrer Sicht besonders schützenswert sind (z.B. Fotos, Adressbuch, Standort oder Kalender).

Die App stellt unterschiedliche Datenschutzerklärungen in der App und im App-Store bereit[128]	
	⊖ Es handelt sich um einen Gesetzesverstoß. ⊖ Der Anbieter ist unseriös. ⊖ Sie wissen nicht, welche Daten zu welchen Zwecken verarbeitet werden.
	⊖ Dies hat nicht zwingend einen Missbrauch Ihrer Daten zur Folge. ⊖ ⊖ Das konkrete Risiko hängt von den tatsächlichen Abweichungen ab.

Verbieten Sie der App alle Zugriffe auf Daten, die aus Ihrer Sicht besonders schützenswert sind (z.B. Fotos, Adressbuch, Standort oder Kalender).

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

[127] Wurde als Rote Linie im Sinne des 4.8.2.2 definiert.
[128] Wurde als Rote Linie im Sinne des 4.8.2.2 definiert.

Die App stellt keine Datenschutzerklärung auf Deutsch bereit	
Anbieter sind gesetzlich verpflichtet, Sie verständlich über die Verarbeitung personenbezogener Daten zu informieren. Dies erfordert grundsätzlich eine deutsche Datenschutzerklärung.	
	⊖ Es handelt sich um einen Gesetzesverstoß. ⊖ Ihre Daten werden bereits verarbeitet, ohne dass Sie sich gegebenenfalls hierüber informieren können.
	⊖ Dies hat nicht zwingend einen Missbrauch Ihrer Daten zur Folge. ⊖ Dies hat nicht zwingend einen Missbrauch Ihrer Daten zur Folge.

Informieren Sie sich durch Medien oder weitere Informationsquellen, ob die Datenverarbeitung der App mit Ihren Datenschutzvorstellungen übereinstimmt.

Verbieten Sie der App alle Zugriffe auf Daten, die aus Ihrer Sicht besonders schützenswert sind (z.B. Fotos, Adressbuch, Standort oder Kalender).

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

Die Datenschutzerklärung verwendet ungenaue Formulierungen	
Die Datenschutzerklärung nutzt Formulierungen wie "z.B.", "in der Regel" oder "grundsätzlich". Sie können sich nicht sicher sein, welche und wie Ihre Daten tatsächlich durch die App verarbeitet werden.	
⊕ Die Datenschutzerklärung kann hierdurch kürzer und übersichtlicher werden.	⊖ Der Anbieter kann die Datenverarbeitungen so verändern, dass diese nicht mehr mit Ihren Vorstellungen übereinstimmen. ⊖ Der Anbieter kommt seiner Aufklärungspflicht nur teilweise nach, sodass Sie sich nicht hinreichend über die Datenverarbeitungen informieren können.
	⊖ Dies ist zulässig, soweit die Änderungen geringfügig sind. Erhebliche Änderungen müssen Ihnen mitgeteilt werden. ⊖

Lesen Sie diese Textabschnitte und entscheiden Sie selbst, ob Sie die App unter diesen Umständen weiterverwenden möchten.

Wenn Sie unsicher sind, ob die Datenverarbeitung Ihren Vorstellungen entspricht, verbieten Sie der App alle Zugriffe auf Daten, die aus Ihrer Sicht besonders schützenswert sind (z.B. Fotos, Adressbuch, Standort oder Kalender).

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

Die App stellt die Datenschutzerklärung erst nach Start der App bereit	
Anbieter sind gesetzlich verpflichtet, Sie bei der Erhebung personenbezogener Daten hierüber zu informieren. Damit Sie selbstbestimmt über Ihre Daten verfügen können, genügt eine Bereitstellung nach Start einer App häufig nicht. Deswegen verlangen auch die App Store-Richtlinien die Bereitstellung bereits vor dem Start, nämlich im App Store.	
	⊖ Ihre Daten werden bereits verarbeitet, ohne dass Sie sich hierüber informieren können. ⊖ Dies verstößt gegen die App Store-Richtlinien.
	⊖ ⊖ Dies hat nicht zwingend einen Missbrauch Ihrer Daten zur Folge.

Verbieten Sie der App alle Zugriffe auf Daten, die aus Ihrer Sicht besonders schützenswert sind (z.B. Fotos, Adressbuch, Standort oder Kalender).

Informieren Sie sich durch Medien, ob die Datenverarbeitung der App mit Ihren Vorstellungen übereinstimmt.

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

Die App klärt nicht ordnungsgemäß über Datenverarbeitungen im Ausland auf	
Wenn und soweit Ihre Daten außerhalb der EU verarbeitet werden, müssen Sie darüber aufgeklärt werden. Hierzu zählt auch, Sie darüber aufzuklären, wie der Anbieter den europäischen Datenschutzstandard außerhalb der EU sicherstellt. Der Standort der Verarbeitung der Daten ist grundsätzlich unabhängig von Ihrem Aufenthaltsort.	
	⊖ Sie wissen gegebenenfalls nicht, ob der Anbieter die sonstigen rechtlichen Voraussetzungen für eine Datenverarbeitung im Ausland beachtet. ⊖ Sie wissen gegebenenfalls nicht, wo Ihre personenbezogenen Daten verarbeitet werden. ⊖ Es handelt sich um einen Gesetzesverstoß.
	⊖ ⊖ ⊖ Dies hat nicht zwingend einen Missbrauch Ihrer Daten zur Folge.

Verbieten Sie der App alle Zugriffe auf Daten, die aus Ihrer Sicht besonders schützenswert sind (z.B. Fotos, Adressbuch, Standort oder Kalender).

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

9.1.3 Datensicherheit

Aspekte die die Absicherung der Daten betreffen. Hierzu zählen Malware oder mangelhafte Verschlüsselungen.

Ihre Login-Daten werden unverschlüsselt übermittelt[129]	
	⊖ Unbefugte Dritte können diese Daten mitlesen und missbrauchen. ⊖ In öffentlichen WLANs können Anmeldeinformationen besonders leicht mitgelesen werden. ⊖ Sollten Sie über mehrere Angebote hinweg die identischen Zugangsdaten (häufig E-Mailadresse und Passwort) nutzen, können Dritte bei diesen Services ungefragten Zugang erhalten.

Loggen Sie sich nicht über die App ein.

Wenn Sie sich trotzdem einloggen möchten, verwenden Sie ein für die App regelmäßig wechselndes Passwort, das Sie bei keinem anderen Nutzerkonto verwenden.

Soweit verfügbar, nutzen Sie fortgeschrittene Authentifizierungsverfahren, wie z.B. Mehrfaktorauthentifizierung.

Wenn Sie bereits Login-Daten über die App übermittelt haben, ändern Sie diese umgehend für andere Nutzerkonten, bei denen Sie dieselben verwenden.

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

Ihre Zahlungsdaten werden unverschlüsselt übermittelt[130]	
	⊖ Unbefugte Dritte können diese Daten mitlesen und missbrauchen, um z.B. Zahlungen zu veranlassen. ⊖ In öffentlichen WLANs können Zahlungsinformationen besonders leicht mitgelesen werden.

Unterlassen Sie die Übermittlung von Zahlungsdaten über die App.

Wenn Sie das nicht möchten, nutzen Sie zur Übermittlung von Zahlungsdaten Online-Bezahlsysteme mit eigener Verschlüsselung.

Wenn Sie bereits Zahlungsdaten über die App übermittelt haben, kontaktieren Sie Ihr Bankinstitut, beobachten Sie Ihre Kontoumsätze und beantragen Sie gegebenenfalls neue Zahlungsmittel.

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

[129] Wurde als Rote Linie im Sinne des 4.8.2.2 definiert.
[130] Wurde als Rote Linie im Sinne des 4.8.2.2 definiert.

Die App enthält Malware[131]	
Die App kann Schadcode ausführen. Beispiele für Malware sind Viren, Trojaner und ähnliches.	
	⊖ Die App kann Ihre Daten missbrauchen. ⊖ Ihr mobiles Gerät oder Daten können beschädigt werden, sodass Sie z.B. nicht mehr auf diese Daten zugreifen können. ⊖ Ihr mobiles Gerät kann dazu genutzt werden, Dritten zu schaden.

Deinstallieren Sie diese App!

Die Verschlüsselung der App ist unsicher	
Die App nutzt eine veraltete oder fehlerhaft implementierte Verschlüsselung.	
	⊖ Unbefugte Dritte können Ihre Daten gegebenenfalls mitlesen und missbrauchen.

Verzichten Sie auf die Übermittlung von Daten, die aus Ihrer Sicht besonders schützenswert sind (die Dritte nicht erfahren sollen, insbesondere in Ihren Nachrichten oder sonstigen Inhalten).

Wenn Sie das nicht möchten, verbieten Sie der App alle Zugriffe auf Daten, die aus Ihrer Sicht besonders schützenswert sind (z.B. Fotos, Adressbuch, Standort oder Kalender).

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

Die App kann sich im Hintergrund unbemerkt aktualisieren	
Die App kann Programmcode ohne Ihre Kenntnis nachladen (Sideloads). Eine Aktualisierung erfolgt nicht über den Play-Store, sondern ohne Hinweis im Hintergrund.	
⊕ Der Anbieter kann die App mit Funktionen erweitern, ohne dass Sie die App updaten müssen.	⊖ Die App kann schädlichen Code (Malware) ohne Ihre Kontrolle nachladen und ausführen. ⊖ Angreifer erhalten eine Möglichkeit, Ihre Daten zu missbrauchen.
⊕ Dies kann z.B. die Integration von Funktionen von Drittanbietern umfassen.	

Deinstallieren Sie die App, es sei denn, Sie vertrauen dem App-Anbieter.

[131] Wurde als Rote Linie im Sinne des 4.8.2.2 definiert.

9.1.4 Identifikation

Aspekte, die die Möglichkeit betreffen, ein Gerät durch den Einsatz digitaler Technologien bei der wiederholten Nutzung und Datenverarbeitung wiederzuerkennen.

Die App erhebt statische Gerätekennungen	
Statische Gerätekennungen sind Kennziffern, die untrennbar mit Ihrem Gerät verbunden sind, z.B. die IMEI, IMSI oder MAC-Adresse. Nicht nur die Kennziffer Ihres mobilen Endgerätes kann betroffen sein, auch die Ihrer verbundenen Geräte und Zubehör (z.B. Kameras oder Docking-Stations). Diese können ohne Spezialkenntnisse nicht vom Nutzer verändert werden. Laut Richtlinien der Betriebssystemhersteller dürfen statische Gerätekennungen nur für Sicherheitszwecke bei sensiblen Services erhoben werden, nicht jedoch für Werbezwecke.	
⊕ Dies kann Ihre Sicherheit erhöhen.	⊖ Ihr Nutzerverhalten kann über mehrere Anwendungen und einen längeren Zeitraum hinweg nachverfolgt werden. ⊖ Eine anonyme Nutzung der App wird erschwert.
⊕ Neben der Authentifizierung durch Ihre Login Daten, kann ein Zugriff über ein berechtigtes Gerät bestätigt werden.	

Wenn Sie nicht möchten, dass statische Gerätekennungen erhoben werden, deinstallieren Sie die App

Die App erhebt eine Vielzahl an Geräteinformationen	
Ihr mobiles Gerät kann hierdurch eindeutig identifiziert werden. Die App erhebt neben den zulässigen Werbe-IDs auch sonstige zur Identifikation dienliche Geräteinformationen (installierte Apps, Spracheinstellungen, Netzbetreiber, Netzempfang, etc.).	
⊕ Dies ermöglicht eine optimierte Funktion oder Darstellung der App.	⊖ Ihr Nutzerverhalten kann über mehrere Anwendungen und einen längeren Zeitraum hinweg nachverfolgt werden.

Entziehen Sie der App bereits vor dem ersten Start alle Zugriffsrechte auf Daten, die aus Ihrer Sicht besonders schützenswert sind (z.B. Fotos, Adressbuch, Standort oder Kalender).

9.1.5 Zugriffe

Aspekte, die das Recht auf Funktionen wie den Standort oder das Adress-buch zuzugreifen, betreffen.

Die App nutzt Standortdaten	
Die App erfasst Ihre Position. Dies kann sowohl durch GPS als auch die Auswertung des Netzwerkempfangs Ihres Mobilfunkbetreibers oder WLAN-Signals geschehen.	
⊕ Standortbezogene Funktionen können personalisiert genutzt werden, ohne dass Sie den Standort extra eingeben müssen.	⊖ Ein dauerhafter Zugriff ermöglicht die Erstellung von Bewegungsprofilen.
⊕ Beispiele sind Navigation und Umgebungssuche oder standortbasierte Werbung.	

Aktivieren Sie nur die standortbezogenen Funktionen Ihres Smartphones, die zum jeweiligen Zeitpunkt benötigt werden (GPS, WLAN, Bluetooth, etc). Wenn Sie das nicht möchten, verbieten Sie der App den Zugriff auf den Standort.

Gewähren Sie der App nur Zugriff auf den genauen Standort (GPS), wenn dies wirklich erforderlich ist. Wenn Sie das nicht möchten, deaktivieren Sie die Standort-Funktion.

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

Die App hat Zugriff auf Ihr Adressbuch	
Die App nutzt Kontaktinformationen, die Sie in Ihrem lokalen Adressbuch gespeichert haben. Dadurch können Ihre Kontakte z.B. auf einer Karte angezeigt oder direkt aus der App heraus kontaktiert werden.	
⊕ Die App kann Kontaktinformationen Ihres Adressbuchs verarbeiten, ohne dass Sie diese extra eingeben müssen.	⊖ Alle gespeicherten Kontaktinformationen können für Zwecke missbraucht werden, die Ihnen unbekannt bleiben.
⊕ Die App kann das Adressbuch aktualisieren oder Kontakte verknüpfen.	⊖ Sie riskieren damit nicht die Übermittlung Ihrer eigenen Daten, sondern auch denjenigen Ihrer Kontakte.

Verbieten Sie der App den Zugriff auf Ihr Adressbuch.

Wenn Sie das nicht möchten, prüfen Sie, ob Sie besonders schützenswerte Kontakte gespeichert haben und löschen Sie diese gegebenenfalls.

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

Die Sprachsteuerung ist dauerhaft im Hintergrund aktiv	
Häufig erfolgt eine Verarbeitung der Gespräche und Umgebungsgeräusche nur sobald ein bestimmtes Signalwort erkannt wird.	
⊕ Sie können Ihr mobiles Endgerät und die App per Sprache steuern.	⊖ Alle Gespräche und Umgebungsgeräusche können dauerhaft aufgenommen und gespeichert werden.

Verbieten Sie den Zugriff der App auf das Mikrofon.

Wenn Sie das nicht möchten, achten Sie darauf, dass die Sprachsteuerung erst nach einem Signalwort aktiviert wird und ansonsten keine Übertragung der aufgenommenen Gespräche und Umgebungsgeräusche stattfindet.

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

9.1.6 Profilbildung

Aspekte, die es ermöglichen, aus den angegebenen und verarbeiteten Informationen ein Profil zu erstellen.

Es wird ein Profil über Sie erstellt	
Der Anbieter, Partner oder integrierte Dritte erstellen ein Profil über Sie. Profile können für unterschiedliche Zwecke genutzt werden; z.B. zur personalisierten Werbung, Individualisierung des Angebots aber auch zur Verbesserung der Sicherheit und Qualität. Profile integrieren in der Regel mehr Informationen (z.B. Informationen über Ihr Nutzungsverhalten oder Geräteinformationen) als solche, die Sie selbst aktiv zur Verfügung gestellt haben. Beachten Sie: Eine Profilbildung ist nicht zwingend mit personalisierter Werbung gleichzusetzen. Wenn Ihre Daten laut Datenschutzerklärung auch für personalisierte Werbung genutzt werden, wird ein entsprechender Fund angezeigt.	
⊕ Dies ermöglicht eine bessere Personalisierung des Angebots.	⊖ Sie können auf Grund Ihres Profils von einzelnen Inhalten und Angeboten ausgeschlossen werden.

Prüfen Sie, ob Sie Möglichkeiten haben, der Profilbildung zu widersprechen.

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

Das über Sie erstellte Profil wird durch öffentliche Informationen über Sie ergänzt	
Öffentliche Informationen über Sie werden Ihrem Profil automatisch hinzugefügt, das durch den Anbieter, einen Partner oder integrierten Dritten über Sie erstellt wurde. Diese öffentlichen Informationen können z.B. aus Ihren Social-Media-Konten, allgemeinen Medienberichten, privaten Webseiten oder öffentlichen Registern entnommen werden. Eine solche Anreicherung, ist gesetzlich grundsätzlich zulässig.	
⊕ Dies ermöglicht eine bessere Personalisierung des Angebots. ⊕ Der Anbieter kann Ihnen die App kostengünstig anbieten.	⊖ Detaillierte Profile bieten ein erhöhtes Missbrauchspotential. ⊖ Sie haben keine Kontrolle darüber, welche Informationen ergänzt werden.

Entscheiden Sie sorgfältig, welche Informationen Sie über andere Apps und Meiden veröffentlichen.

Wenn Sie wissen möchten, welche öffentlichen Informationen die App in Ihrem Profil ergänzt, verlangen Sie eine Auskunft bei dem App-Anbieter.

Werden öffentliche Informationen zu Ihrem Profil hinzugefügt und der App-Anbieter kann hierfür keinen hinreichenden Grund nennen, beantragen Sie bei diesem die Löschung dieser Daten.

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

9.1.7 Werbung

Aspekte bei denen Werbung durch Drittanbieter ausgespielt oder aufgrund der Datenverarbeitung auf eine Person oder ein Gerät zugeschnitten wird.

Die App integriert Werbenetzwerke	
Die App zeigt Werbung externer Werbenetzwerke an. Somit erhalten neben dem Anbieter auch Dritte Daten über Sie. Beachten Sie: Die Integration von Werbenetzwerken geht nicht zwingend mit personalisierter Werbung einher. Wenn Ihre Daten laut Datenschutzerklärung auch für personalisierte Werbung genutzt werden, wird ein entsprechender Fund angezeigt.	
⊕ Der Anbieter kann Ihnen die App kostengünstig anbieten.	⊖ Ihr Nutzungsverhalten kann über mehrere Anwendungen und einen längeren Zeitraum hinweg nachverfolgt werden. ⊖ Diese Drittanbieter haben grundsätzlich die gleichen Zugriffsrechte die die App, sodass neben den Nutzungsdaten auch Ihre sonstigen daten übermittelt werden können.
⊕ Durch die Nutzung von Werbenetzwerken muss der Anbieter z.B. keine kostenintensive, eigene Werbekundenverwaltung finanzieren.	

Wenn Sie nicht möchten, dass Werbenetzwerke Daten über Sie erhalten, deinstallieren Sie die App!

Ihre Daten werden für personalisierte Werbung genutzt	
⊕ Der Anbieter kann Ihnen die App kostengünstig anbieten.	⊖ Voraussetzung für personalisierte Werbung ist eine Profilbildung.
⊕ Der Anbieter kann über verkaufte Werbeplätze Geld verdienen und hierdurch die App finanzieren.	

Falls Sie grundsätzlich keine personalisierte Werbung erhalten möchten, deaktivieren Sie diese.

Ändern Sie auch regelmäßig Ihre dynamischen Werbe-IDs, um einer Nachverfolgung und Profilbildung entgegenzuwirken.

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

9.1.8 Übertragung an Dritte

Aspekte, die die Weitergabe von verarbeiteten Daten an Dritte betreffen.

Die App übermittelt Daten an Dritte	
In eine solche Übermittlung müssen Sie explizit einwilligen. Eine Einwilligung ist nicht erforderlich, wenn es ausnahmsweise eine gesetzlich Sondervorschrift (z.B. zur Rechtsdurchsetzung und Strafverfolgung) gibt.	
⊕ Kann Rechtsdurchsetzung erleichtern.	⊖ Dritte können Ihre Daten zu unerwarteten eigenen Zwecken verarbeiten.

Prüfen Sie, ob Sie in eine solche Übermittlung eingewilligt haben. Wenn Sie mit einer Übermittlung nicht mehr einverstanden sind, können Sie die Einwilligung jederzeit widerrufen.

Verbieten Sie der App Zugriffe auf Daten ein, die aus Ihrer Sicht besonders schützenswert sind (z.B. Fotos, Adressbuch, Standort oder Kalender).

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

Ihre Daten werden durch Dienstleister verarbeitet	
Dienstleister ist jedes Unternehmen, welches den Anbieter beim Angebot der Services unterstützt. Dieses Verhältnis nennt das Gesetz Auftragsverarbeitung. Das Gesetz erleichtert den Einsatz von Auftragsverarbeitern, stellt dafür aber strenge Anforderungen, um die personenbezogenen Daten vor Missbrauch zu schützen.	
⊕ Der Anbieter kann Ihnen die App kostengünstig anbieten. ⊕ Dies kann die Qualität und Sicherheit der App erhöhen.	⊖ Sie können die Weitergabe Ihrer Daten nur begrenzt kontrollieren. ⊖ Sie wissen nicht, ob diese Dienstleister den gleichen Standard wie der Anbieter im Umgang mit Ihren Daten haben.
⊕ Dies erfolgt z.B. durch den Einsatz spezialisiert Dienstleister oder erprobter Plugins und Infrastruktur.	⊖ Ihre personenbezogenen Daten werden grundsätzlich vom Anbieter und nach dessen Vorgaben verarbeitet. Jedoch ist es möglich, dass ein unseriöser Dienstleister Ihre Daten ohne Ihr Wissen und das Wissen des App-Anbieters weiterverarbeitet. ⊖ Ihre personenbezogenen Daten werden grundsätzlich vom Anbieter und nach dessen Vorgaben verarbeitet. Jedoch ist es möglich, dass ein unseriöser Dienstleister Ihre Daten ohne Ihr Wissen und das Wissen des App-Anbieters weiterverarbeitet.

Verbieten Sie der App Zugriffe auf Daten ein, die aus Ihrer Sicht besonders schützenswert sind (z.B. Fotos, Adressbuch, Standort oder Kalender) oder deinstallieren Sie die App.

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

Ihre Daten werden über die App veröffentlicht	
Sie können über die App Inhalte erstellen und mit allen Nutzern der App bzw. der Öffentlichkeit teilen. Diese Daten können z.B. auch von Dritten genutzt werden, um Ihre Profilinformationen anzureichern.	
⊕ Sie können Ihre Beiträge leichter einem breiten Publikum zugänglich machen.	⊖ Sind Ihre Beiträge öffentlich im Internet verfügbar, haben Sie keine Kontrolle, wer auf diese zugreift und zu welchen Zwecken diese Daten weiterverarbeitet werden.

Überlegen Sie vor Veröffentlichung Ihrer Beiträge, ob diese Inhalte wirklich für eine unbekannte Öffentlichkeit bestimmt sind.

Schränken Sie, wenn möglich, den Zugriff auf Ihre Beiträge ein.

Die App ermöglicht einer Vielzahl von Drittanbietern Zugriff auf Ihre Nutzungsdaten	
Ein Drittanbieter ist jede datenverarbeitende Stelle, die personenbezogene Daten von Ihnen erhält, aber nicht selbst Anbieter der App ist. Hierzu zählen neben Analyse-, Tracking- und Werbetools auch Social-Media-Services. Nutzungsdaten sind insbesondere Merkmale zur Identifikation des Nutzers, Angaben über Beginn und Ende der jeweiligen Nutzung sowie Angaben über die vom Nutzer in Anspruch genommene Internetdienste.	
⊕ Die App kann Funktionen der Drittanbieter in gewohnter Form bereitstellen ⊕ Der Anbieter kann Ihnen die App kostengünstig anbieten.	⊖ Ihr Nutzungsverhalten kann über mehrere Anwendungen und einen längeren Zeitraum hinweg nachverfolgt werden. ⊖ Diese Drittanbieter haben grundsätzlich die gleichen Zugriffsrechte wie die App, sodass neben den Nutzungsdaten auch Ihre sonstigen Daten übermitteln werden können.

Sollten Drittanbieter integriert sein, denen Sie keine Daten übermitteln möchten, prüfen Sie in der App, ob Sie einer solchen Datenübermittlung widersprechen können.

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

Ihre Daten werden in der Unternehmensgruppe geteilt	
Eine Unternehmensgruppe besteht in der Regel aus einem leitenden Unternehmen sowie weiteren Tochterunternehmen, die von diesem abhängig sind	
⊕ Sie können in der App auch auf andere Services der Unternehmensgruppe zugreifen. ⊕ Dies kann die Qualität und Sicherheit der App erhöhen.	⊖ Sie können die Weitergabe Ihrer Daten nur begrenzt kontrollieren.
⊕ ⊕ Innerhalb der Unternehmensgruppe können sich spezialisierte Abteilungen gezielt mit auftretenden Themen beschäftigen.	⊖ Ihre personenbezogenen Daten müssen jedoch grundsätzlich nach einheitlichen Standards verarbeitet werden.

Achten Sie darauf, welche Ihrer personenbezogenen Daten laut Datenschutzerklärung in der Unternehmensgruppe geteilt werden können.

Verbieten Sie der App Zugriffe auf Daten, die aus Ihrer Sicht besonders schützenswert sind (z.B. Fotos, Adressbuch, Standort oder Kalender).

Sollte(n) diese Empfehlung(en) nicht ausreichen, deinstallieren Sie die App!

9.2 Befragung 1

9.2.1 Filterfragen

Vor Beginn des Fragebogens wurden den Befragten drei Filterfragen gestellt.

1. Nutzen Sie ein Smartphone? Ja Nein
2. In welchem Jahr sind Sie geboren? ____
3. Sind Sie weiblich oder männlich? Weiblich männlich

Wurde die erste Filterfrage mit „nein“ beantwortet, wurde die Umfrage beendet. Teilnehmerinnen und Teilnehmer, die ein Smartphone nutzen, gelangen zu den weiteren Fragen, die im Folgenden dargestellt sind. Zusätzlich wurde zur Sicherstellung der Repräsentativität das Bundesland, in dem die Teilnehmerinnen und Teilnehmer ihren ersten Wohnort haben, abgefragt.

9.2.2 Fragebogen

Sehr geehrte Teilnehmerin, sehr geehrter Teilnehmer,

vielen Dank für Ihre Teilnahme an unserer Umfrage!

Diese Umfrage wurde von einer Forschungsgruppe an der Quadriga Hochschule Berlin gestaltet, die in einem Forschungsprojekt den Schutz von personenbezogenen Daten bei Smartphones untersucht.

Dieses vom Bundesministerium für Bildung und Forschung (BMBF) geförderte Projekt zielt darauf ab, eine Smartphone-App zu entwickeln, die Verbraucherinnen und Verbraucher darin unterstützt, sich selbst besser vor möglichen Gefahren bei der Nutzung von Smartphone-Apps zu schützen. Die Ergebnisse der Befragung und Ihre Antworten sind von hoher Bedeutung und fließen in die Entwicklung der App ein.

Für die Umfrage sollten Sie sich ca. 15 Minuten Zeit nehmen. In der Befragung geht es um Ihre persönliche Meinung.

Die Umfrage wird ausschließlich von den Mitarbeiterinnen und Mitarbeitern des Forschungsprojekts ausgewertet.

Ihre Antworten sind hierbei vollkommen anonym und es sind keine Rückschlüsse auf Ihre Person möglich.

Falls Sie Fragen oder Anmerkungen zu dieser Studie haben, senden Sie uns eine

E-Mail.

Zuerst erhalten Sie einige allgemeine Fragen zu Ihrer Nutzung von Smartphones und Smartphone-Apps.
Die Fragen beziehen sich auf Ihr aktuelles Smartphone. Falls Sie mehrere Smartphones besitzen, beantworten Sie die Fragen bitte für das von Ihnen am häufigsten genutzte Smartphone:

Q1a
Wie viele Apps haben Sie auf Ihrem aktuellen Smartphone selbst installiert, d.h. aus dem Appstore geladen?

o Keine (0)
o zwischen 1 und 5
o zwischen 6 und 10
o zwischen 11 und 20
o zwischen 21 und 30
o mehr als 30
o weiß nicht

Falls eine oder mehrere Apps installiert oder Antwort „weiß nicht" → Folgefrage Q1b, sonst weiter mit Q2

Q1b
Haben Sie eine oder mehrere Apps gekauft, d.h. für diese Geld bezahlt?
o Ja o Nein

Q2
Welches Betriebssystem hat Ihr aktuelles Smartphone?
o Android
o iOS
o Windows
o Anderes
o weiß ich nicht

Q3
Unter *Datenschutz bei der Verwendung von Smartphones* verstehen wir im Folgenden den *Schutz Ihrer personenbezogenen Daten bei der Nutzung von vor- oder selbst installierten Smartphone-Apps*.

Wie hoch schätzen Sie Ihren Wissensstand zum Thema Schutz personenbezogener Daten bei der Verwendung von Smartphone-Apps ein?

o	o	o	o	o
sehr gering	gering	mittelmäßig	hoch	sehr hoch

Die nächsten vier Fragen behandeln konkrete Aspekte zum Thema Datenschutz bei der Nutzung von Smartphone-Apps. Zu jeder Frage werden Ihnen vier Antwortmöglichkeiten angeboten, von denen nur eine Antwort korrekt ist.

Die Antwortmöglichkeiten der Fragen Q4 bis Q7 wurden in zufälliger Reihenfolge angezeigt. Zu jeder Frage erhielten die Teilnehmer folgende Ausfüllanweisung:

Nur eine Antwortmöglichkeit ist richtig.
Bitte wählen Sie die Antwortmöglichkeit aus, die Ihrer Meinung nach die richtige ist.

Q4
Welche Einstellungen bei Smartphones verbessern den Datenschutz?
o Die Voreinstellungen sind schon datenschutzgerecht, sodass keine Verbesserungen nötig sind.
o Ortungsdienste und Schnittstellen, die gerade nicht gebraucht werden (z.B. GPS, WLAN, Bluetooth), sollten deaktiviert werden.
o Es sollte immer ein aktueller Virenscanner installiert sein, dann werden Daten nicht ungewollt verschickt.
o Die aktuelle Version des Betriebssystems garantiert, dass Daten nicht heimlich verarbeitet werden.

Q5

Muss man das "Kleingedruckte" in den Datenschutzbestimmungen einer App immer lesen?

o Ja. Wer auf "gelesen", "einverstanden" oder Ähnliches klickt, ohne die Texte gelesen zu haben, macht sich strafbar.

o Nein, nicht bei kostenlosen Apps. Egal, was in den Texten steht: Nachteile sind nicht zu befürchten, wenn die App nichts kostet.

o Nein, man muss nicht. Aber man sollte: In den Texten können sich Datenverarbeitungen verbergen, die auf den ersten Blick nicht zu erkennen sind.

o Nein. Die Texte sind nur zur Information. Alles Wichtige muss man unterschreiben.

Q6

Für die Nutzung einer App muss man dem Anbieter eine Einwilligung für die Datenverarbeitung erteilen. Was trifft auf eine Einwilligung im Sinne der Datenschutzgesetze zu?

o Man muss einer Datenschutzerklärung explizit zustimmen und kann diese im Nachhinein nicht widerrufen.

o Man muss einer Datenschutzerklärung explizit zustimmen und die Einwilligung gilt dann unabhängig von der Zweckmäßigkeit für alle Datenverarbeitungen der App.

o Die Einwilligung ist freiwillig und kann auch im Nachhinein widerrufen werden.

o Die Einwilligung ist auch gültig, wenn sie nicht ausdrücklich erteilt wurde und beginnt mit dem Aufruf und der ersten Nutzung der App.

Q7

Sie installieren eine ausländische App, die sich an deutsche Nutzer richtet. In welcher Sprache muss die zugehörige Datenschutzerklärung der App verfasst sein?

o In einer der Amtssprachen der EU, zum Beispiel Deutsch, Englisch oder Französisch

o In der Sprache des Anbieters

o In deutscher Sprache

o Hierzu gibt es keine Regeln

Q8

Die nächste Frage bezieht sich auf Ihr eigenes Verhalten beim Surfen im Internet und bei der Verwendung Ihres Smartphones:

Haben Sie in der Vergangenheit bereits konkrete Maßnahmen zum Schutz Ihrer personenbezogenen Daten im Internet oder im Smartphone getroffen?

Mehrfachnennung möglich

Ja,
- o ich habe von meinen Betroffenenrechten Gebrauch gemacht (Widerruf, Auskunft, Berichtigung, Sperrung).
- o ich habe die MAC-Adresse oder sonstige Gerätekennungen meines PCs, Laptops oder Smartphones verändert.
- o ich nutze Pseudonyme, z.B. in sozialen Netzwerken wie Facebook oder Google+.
- o ich habe Markierungen und Verlinkungen durch Dritte entfernt (z.B. Bildmarkierungen in sozialen Netzwerken).
- o ich habe mich durch Lesen der Datenschutzerklärung über die Datenverarbeitung einer Webseite informiert.
- o ich habe mich durch Lesen der Datenschutzerklärung über die Datenverarbeitung einer App informiert.
- o ich habe Berechtigungen oder Funktionszugriffe (z.B. Standort, Adressbuch oder Kamera) von Apps entzogen.
- o ich habe Apps deinstalliert.
- o Sonstiges: _____

o Nein

Q9
Wie bereits in der Einleitung erwähnt, geht es in unserem Projekt darum, eine Smartphone-App zu entwickeln, die Verbraucherinnen und Verbraucher darin unterstützt, sich selbst besser vor möglichen Gefahren bei der Nutzung von Smartphone-Apps zu schützen.

Stellen Sie sich nun bitte vor, dass eine solche Selbstdatenschutz-App bereits existiert und mehrere Funktionen bietet, um den Schutz Ihrer personenbezogenen Daten bei der Verwendung vor- und selbst installierter Smartphone-Apps zu erleichtern.
Antwortoptionen wurden über zwei Seiten in Tabellenform randomisiert. Auf der ersten Seite wurden zufällig fünf Items angeordnet, auf der nächsten Seite der Befragung zufällig die übrigen vier Items.

Wie wichtig oder unwichtig wären Ihnen die folgenden Funktionen bei einer Selbstdatenschutz-App?
Die App soll…

	Sehr unwichtig	Unwichtig	Weder wichtig, noch unwichtig	Wichtig	Sehr wichtig
… mich über die Datenverarbeitung der einzelnen Apps, die ich bereits auf meinem Smartphone installiert habe, informieren.	o	o	o	o	o
… mir eine Übersicht der Datenverarbeitungen aller installierten Apps bieten.	o	o	o	o	o
… mir besonders kritische Aspekte automatisch anzeigen.	o	o	o	o	o
… mich vor Installation neuer Apps über deren Datenverarbeitung informieren.	o	o	o	o	o
… mich informieren, wenn ein App-Anbieter seine Datenschutzerklärung und die Datenverarbeitung verändert.	o	o	o	o	o
… mir Handlungsempfehlungen geben, wie ich den Schutz meiner Daten verbessern kann.	o	o	o	o	o
… mir anzeigen, wie andere Nutzer oder Institutionen (z. B. Ministerien oder Datenschützer) die von mir installierten Apps bewerten.	o	o	o	o	o
… mir Alternativen zu meinen installierten Apps aufzeigen.	o	o	o	o	o
… Datenschutzerklärungen von Apps für mich verständlicher machen.	o	o	o	o	o

Q10
Welche weiteren Funktionen sind aus Ihrer Sicht für eine solche Datenschutz-App wichtig?
_____

Q11
Funktionen der App:

- Informiert über die Datenverarbeitung der einzelnen Apps, die bereits auf dem Smartphone installiert sind.
- Bietet eine Übersicht der Datenverarbeitungen aller installierten Apps.
- Zeigt besonders kritische Aspekte automatisch an.
- Informiert vor der Installation neuer Apps über deren Datenverarbeitung.
- Informiert, wenn ein App-Anbieter seine Datenschutzerklärung und die Datenverarbeitung verändert.
- Gibt Handlungsempfehlungen, wie man den Schutz der eigenen Daten verbessern kann.
- Zeigt an, wie andere Nutzer oder Institutionen (zum Beispiel Ministerien oder Datenschützer) die bereits installierten Apps bewerten.
- Zeigt Alternativen zu den bereits installierten Apps auf.
- Macht Datenschutzerklärungen von Apps verständlicher.

Stellen Sie sich vor, die App hätte alle oben genannten Funktionen: Würden Sie diese App installieren?

o Ja o Nein o weiß ich nicht

Q12
Wie viel würden Sie für eine App mit den oben genannten Funktionen bezahlen?
Wie viel? ___ Euro
(Bitte geben Sie den Wert in Euro an; falls Sie kein Geld für diese App zahlen würden, geben Sie bitte 0 an)

Q13 a
Falls Q11 == „ja“
Würden Sie zusätzlich zu dieser App eine Webseite nutzen, die Ihnen weiterführende Informationen und Neuigkeiten zum Thema Schutz personenbezogener Daten bei der Verwendung von Smartphone-Apps liefert und Statistiken zur Datenverarbeitung beliebter Apps aufbereitet?
o Ja o Nein o weiß ich nicht

Q13 b
Falls Q11 == „nein“
Würden Sie anstatt dessen eine Webseite nutzen, die Ihnen weiterführende Informationen und Neuigkeiten zum Thema Schutz personenbezogener Daten bei der Verwendung von Smartphone-Apps liefert und Statistiken zur Datenverarbeitung beliebter Apps aufbereitet?
o Ja o Nein o weiß ich nicht

Q13 c
Falls Q11 == „weiß nicht“
Würden Sie eine Webseite nutzen, die Ihnen weiterführende Informationen und Neuigkeiten zum Thema Schutz personenbezogener Daten bei der Verwendung von Smartphone-Apps liefert und Statistiken zur Datenverarbeitung beliebter Apps aufbereitet?
o Ja o Nein o weiß ich nicht

Q14 bis Q26
Im Folgenden erhalten Sie 13 Hinweise, welche unsere Datenschutz-App, die wir im Rahmen des Forschungsprojekts entwickeln, geben könnte. Diese Hinweise beziehen sich auf die Datenverarbeitung und die Datenschutzerklärungen von Apps, die Sie zurzeit aus dem App-Store laden können. Alle Hinweise sind realistisch und könnten somit auch Ihr Smartphone betreffen.

Ein Beispiel könnte der folgende sehr kritische Hinweis sein:

Ihre Zahlungsdaten werden unverschlüsselt übermittelt[132]	
	⊖ Unbefugte Dritte können entsprechende Daten mitlesen und missbrauchen. ⊖ In öffentlichen WLANs können Zahlungsinformationen leicht mitgeschnitten werden.

Ein weiteres Beispiel ist der folgende weniger kritische Hinweis:

Die App verfügt über Internetzugriff	
⊕ Sie können aktuelle Inhalte nutzen. ⊕ Sie können weiterführende Komfortfunktionen aktivieren (z.B. Ticketbuchungen).	⊖ Die App kann alle Daten, auf die sie zugreift, auch versenden.

Bitte lesen Sie die 13 Hinweise genau und bewerten diese auf folgender Skala:

Falls Ihr Smartphone von dieser Datenverarbeitung betroffen ist oder eine ähnlich ungenügende Datenschutzerklärung vorlegt, wie sehr interessiert Sie dann ein Hinweis darüber?

Interessiert mich überhaupt nicht o o o o o o o interessiert mich sehr

[132] Wurde als Rote Linie im Sinne des 4.8.2.2 definiert.

Aus den insgesamt 26 Infoblöcken wurden 13 zufällig ausgewählt und von den Teilnehmern bewertet.

Die App hat Zugriff auf Ihr Adressbuch	
⊕ Dies kann Ihren Komfort erhöhen, weil die App Kontaktinformationen Ihres Adressbuches verarbeiten kann (z.B. Messaging und Navigation).	⊖ Alle gespeicherten Kontaktinformationen können für Ihnen unbekannte Zwecke missbraucht werden.

Die App ermöglicht einer Vielzahl von Drittanbietern Zugriff auf Ihre Nutzungsdaten	
⊕ Dies kann den Komfort erhöhen, da die App Funktionen in gewohnter Form bereitstellen kann (z.B. Zahlungsdienste oder Chats). ⊕ Dies ermöglicht dem Anbieter, Ihnen die App kostengünstig anzubieten.	⊖ Drittanbieter können Nutzungsprofile anlegen. ⊖ Dies ermöglicht, Ihr Nutzungsverhalten über mehrere Apps hinweg zu verfolgen.

Die App stellt die Datenschutzerklärung erst nach Start der App bereit	
⊕ Ihre Persönlichkeitsrechte können trotzdem hinreichend geschützt sein.	⊖ Ihre Daten werden gegebenenfalls bereits verarbeitet, bevor Sie sich hierüber eine Meinung bilden können. ⊖ Dies verstößt gegen die App Store-Richtlinien.

Die App stellt keine Datenschutzerklärung bereit	
	⊖ Es handelt sich um einen Gesetzesverstoß.

Die Datenschutzerklärung verwendet ungenaue Formulierungen	
⊕ Die Datenschutzerklärung kann hierdurch kürzer und übersichtlicher werden.	⊖ Der Anbieter kann die Datenverarbeitungen so verändern, dass sie nicht mehr mit Ihren Vorstellungen übereinstimmen. ⊖ Der Anbieter kommt seiner Aufklärungspflicht nur teilweise nach. Sie können sich keine genaue Vorstellung über die Datenverarbeitungen bilden.

Die Datenschutzerklärung beinhaltet einen Änderungsvorbehalt ohne Information der Nutzer	
	⊖ Es handelt sich um einen Gesetzesverstoß. ⊖ Der Anbieter kann die Datenverarbeitungen so verändern, dass sie nicht mehr mit Ihren Vorstellungen übereinstimmen. ⊖ Sie werden über diese Veränderungen nicht informiert.

Die App stellt keine Datenschutzerklärung auf deutsch bereit	
⊕ Ihre Persönlichkeitsrechte können trotzdem hinreichend geschützt sein.	⊖ Ihre Daten werden bereits verarbeitet, ohne dass Sie sich hierüber eine eigene Meinung bilden können. ⊖ Es handelt sich um einen Gesetzesverstoß.

Die App benennt keine Kontaktmöglichkeit für datenschutzrechtliche Anliegen	
⊕ Ihre Persönlichkeitsrechte können trotzdem hinreichend geschützt sein.	⊖ Ihre Anfragen können verloren gehen. ⊖ Es handelt sich um einen Gesetzesverstoß.

Die App klärt nicht ordnungsgemäß über Datenverarbeitungen außerhalb der EU auf	
⊕ Ihre Persönlichkeitsrechte können trotzdem hinreichend geschützt sein.	⊖ Ihre personenbezogenen Daten können außerhalb der EU schlechter geschützt sein. ⊖ Sie wissen nicht, wo Ihre personenbezogenen Daten verarbeitet werden.

Die App verarbeitet Daten, die für die Funktion der App nicht erforderlich sind	
⊕ Dies ermöglicht es, die App um Komfortfunktionen zu erweitern (z.B. Kalendereinträge, Kartendarstellung).	⊖ Es werden unnötige Daten erhoben. ⊖ Ihre Daten können zu detaillierten Profilen zusammengeführt werden. ⊖ Ihre Daten können missbraucht werden.

Die App verarbeitet Daten, die ausdrücklich ausgeschlossen wurden	
	⊖ Es handelt sich um eine Täuschung und einen Gesetzesverstoß. ⊖ Der Anbieter ist unseriös. ⊖ Es besteht eine erhöhte Missbrauchsgefahr für Ihre Daten.

Die App erhebt eine Vielzahl an Geräteinformationen, sodass Ihr Gerät eindeutig identifiziert werden kann	
⊕ Dies ermöglicht eine auf Ihr Gerät optimierte Funktion oder Darstellung der App.	⊖ Dies ermöglicht, das Nutzerverhalten über mehrere Anwendungen und einen längeren Zeitraum hinweg zu verfolgen. ⊖ Die App erhebt mehr Informationen als notwendig.

Es erfolgt eine Profilbildung für Drittzwecke	
⊕ Dies ermöglicht eine bessere Personalisierung des Service. ⊕ Dies ermöglicht dem Anbieter, Ihnen die App kostengünstig anzubieten.	⊖ Sie können nicht kontrollieren, welche Informationen über Sie in Profilen hinterlegt werden und wie häufig diese mit welchen Daten angereichert werden.

Profile werden durch öffentliche Drittinformationen über Sie angereichert	
⊕ Dies ermöglicht eine bessere Personalisierung des Service. ⊕ Dies ermöglicht dem Anbieter, Ihnen die App kostengünstig anzubieten.	⊖ Detaillierte Profile bieten ein erhöhtes Missbrauchspotential. ⊖ Sie haben keine Kontrolle darüber, welche Informationen eingekauft und in welcher Weise in das Profil eingespeist werden.

Die App erhebt statische Gerätekennungen	
⊕ Dies kann die Sicherheit von sicherheitsrelevanten, kritischen Apps verbessern.	⊖ Dies ermöglicht, Ihr Nutzerverhalten über mehrere Anwendungen und einen längeren Zeitraum hinweg nachzuverfolgen. ⊖ Dies verhindert Ihre Anonymität.

Ihre Daten werden in der Unternehmensgruppe geteilt	
⊕ Dies vereinfacht die Erstellung eines Nutzerprofils, das in mehreren Apps des gleichen Anbieters genutzt werden kann und Zugriff auf weiterführende Funktionen bietet (z.B. Verknüpfung mehrerer Services eines App-Anbieters miteinander).	⊖ Sie können die Weitergabe von Informationen über sich nicht kontrollieren. ⊖ Nutzerprofile/-daten sind nur noch schwer löschbar und Sie können die Löschung nicht nachvollziehen.

Die App nutzt Standortdaten	
⊕ Dies ermöglicht, standortbezogene Services anzubieten (Navigation, Streckenverfolgung (Jogging), Umgebungssuche). ⊕ Sie müssen Ihren Standort nicht gesondert eintippen.	⊖ Ein dauerhafter Zugriff ermöglicht die Erstellung von Bewegungsprofilen.

Ihre Daten werden durch Dritte und Partner für individualisierte Werbung genutzt	
⊕ Der Anbieter kann Ihnen die App kostengünstig anbieten. ⊕ Hierdurch kann personalisierte Werbung angezeigt werden.	⊖ Zur personalisierten Werbung muss das Werbenetzwerk Sie identifizieren. ⊖ Dies ermöglicht, Ihr Nutzungsverhalten über mehrere Anwendungen und einen längeren Zeitraum hinweg nachzuverfolgen. ⊖ Die App kann Werbung aus Suchverläufen und weiteren Konten wie Shopping-Portalen oder Social-Media anzeigen.

Ihre Login-Daten werden unverschlüsselt übermittelt	
	⊖ Unbefugte Dritte können entsprechende Daten mitlesen und missbrauchen. ⊖ In öffentlichen WLANs können Anmeldeinformationen leicht mitgeschnitten werden.

Die App enthält Malware	
	⊖ Die App kann Schadcode ausführen und z.B. Zugangsdaten oder Bankdaten abgreifen. ⊖ Ihr mobiles Gerät kann unbrauchbar gemacht werden, indem alle darauf befindlichen Daten verschlüsselt werden.

Die Herkunft der App ist unbekannt	
⊕ Dies ermöglicht die Verwendung von Apps, die nicht in einem App Store gelistet werden.	⊖ Die App könnte Malware enthalten. ⊖ Die App könnte eine Raubkopie sein.

Die Verschlüsselung der App ist unsicher implementiert	
	⊖ Dies erleichtert es Dritten, unbefugt Ihre Kommunikation mitzulesen und gegebenenfalls zu missbrauchen.

Ihre Daten werden durch Dienstleister verarbeitet	
⊕ Dies ermöglicht dem Anbieter die App kostengünstig anzubieten. ⊕ Die Einbindung spezialisierter Dienstleister kann die Sicherheit erhöhen, weil erprobte Plugins und Infrastruktur genutzt werden.	⊖ Der Nutzer hat nur sehr begrenzte Kontrolle, wohin seine Daten fließen und ob alle Dienstleister das gleiche (hohe) Verständnis bzgl. Datenschutz und Datensicherheit haben.

Änderungen der App-Technik im Hintergrund möglich	
⊕ Der Anbieter kann die App mit Funktionen erweitern, ohne dass der Nutzer die App updaten muss.	⊖ Sie haben keine Kontrolle darüber, dass die App Code aus dem Internet nachladen und ausführen kann (z.B. Schadcode/Malware). ⊖ Angreifer erhalten eine Möglichkeit, Ihr Gerät an- und Ihre Daten abzugreifen.

Die App integriert Werbenetzwerke	
⊕ Der Anbieter kann Ihnen die App kostengünstig anbieten. ⊕ Es kann personalisierte Werbung angezeigt werden.	⊖ Dies ermöglicht, das Nutzungsverhalten über mehrere Anwendungen und einen längeren Zeitraum hinweg nachzuverfolgen. ⊖ Die App kann Werbung aus Suchverläufen, weiteren Konten wie Shopping-Portalen oder Social-Media anzeigen.

Die App übermittelt Daten zu Werbezwecken an Dritte	
	⊖ Sie können nicht abschätzen, wer Daten erhält. ⊖ Käufer der Daten können wiederum Profile über Sie mit weiteren Informationen anreichern und unkontrolliert nutzen.

Abschließend haben wir noch ein paar Fragen zu Ihrer Person.

Q27

Wenn man alle Einkünfte zusammennimmt: Wie hoch ist Ihr monatliches Einkommen?

Bitte geben Sie den monatlichen Netto-Betrag an, also nach Abzug von Steuern und Sozialabgaben. Regelmäßige Zahlungen wie Renten, Wohngeld, Kindergeld, BAföG, Unterhaltszahlungen usw. rechnen Sie bitte dazu!

o Weniger als 1000 €
o 1000 bis 1999 €
o 2000 bis 2999 €
o 3000 bis 3999 €
o Mehr als 4000 €
o keine Angabe

Q28

Wie ist Ihre derzeitige berufliche Situation?

o Voll erwerbstätig
o In Teilzeitbeschäftigung
o In betrieblicher Ausbildung / Lehre
o Schüler/in, Student/in
o Geringfügig oder unregelmäßig erwerbstätig
o Rentner/in, Pensionär/in
o Nicht erwerbstätig
o Andere: _____

Q29

Was ist Ihr höchster Bildungsabschluss?

o Hauptschulabschluss
o Mittlerer Schulabschluss (z.B. Realschulabschluss)
o Fachhochschulreife (Abschluss einer Fachoberschule)
o Abitur (Hochschulreife)
o Bachelor
o Master / Diplom
o Promotion
o Habilitation
o Andere: _____

Q30

Welche Staatsangehörigkeit haben Sie?

_____

Q31

Welche Sprache wird in Ihrem Haushalt hauptsächlich gesprochen?

_____

Vielen Dank für Ihre Teilnahme!

Impressum:
PrivacyGuard Projekt
Quadriga Hochschule Berlin
Werderscher Markt 13
10117 Berlin

Falls Sie Fragen zum PrivacyGuard Projekt haben und direkt Kontakt mit der Forschergruppe an der Quadriga Hochschule aufnehmen möchten, senden Sie eine E-Mail.

9.2.3 Beschreibung der Stichprobe

Im Rahmen der ersten Online-Umfrage wurden insgesamt 1.000 Menschen mit Smartphone befragt. Ungewichtet weist die Stichprobe folgende Merkmale auf:

9.2.3.1 Alter

Der jüngsten Alterskohorte (14-19 Jahre) gehören 9% der Befragten an. 17% der Umfrageteilnehmer sind zwischen 20 und 29 Jahren, 16% zwischen 30 und 39 Jahren, 20% zwischen 40 und 49 Jahren, 17% zwischen 50 und 59 Jahren, 10% zwischen 60 und 69 Jahren und 11% zwischen 70 und 79 Jahren. Weitere 1% der Teilnehmer sind über 80 Jahre alt.

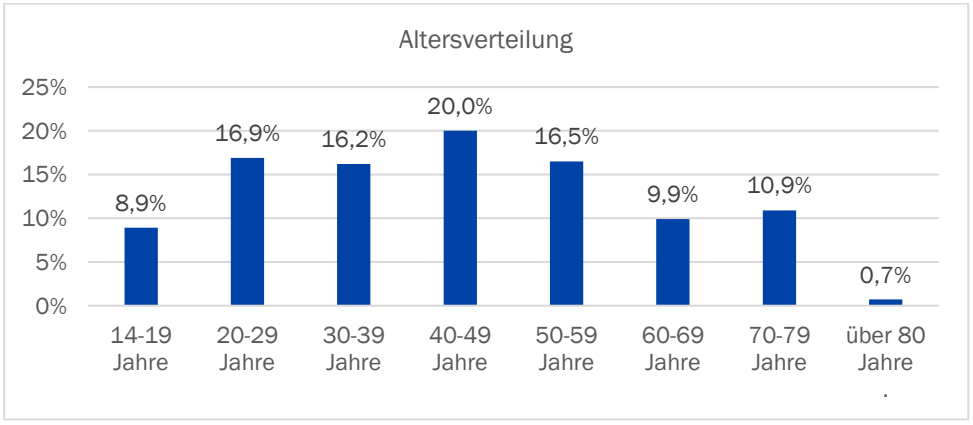


Abbildung 92: Alter der Befragten – Befragung 1

9.2.3.2 Geschlecht

50,5% der Umfrageteilnehmer sind männlich, 49,5% weiblich.

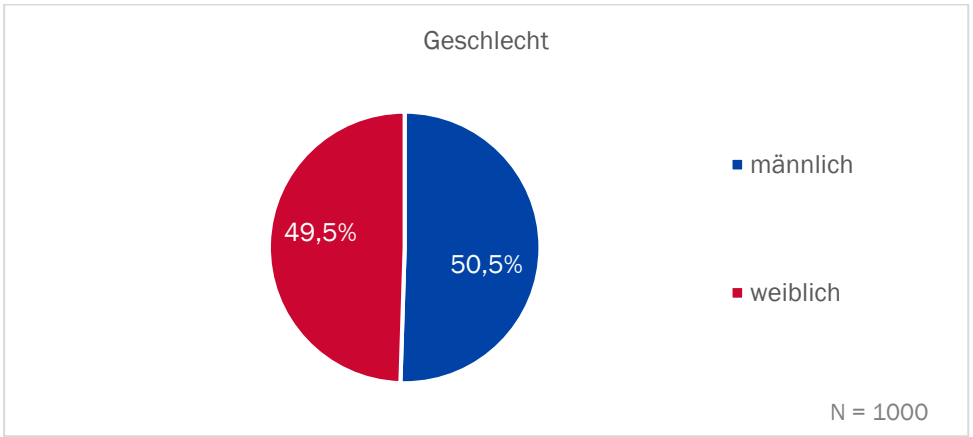


Abbildung 93: Geschlecht der Befragten

9.2.3.3 Nettoeinkommen

Hinsichtlich des monatlichen Nettoeinkommens gaben 21% der Teilnehmer an weniger als 1000€ zur Verfügung zu haben. 36% verfügen über ein Nettoeinkommen zwischen 1000 und 1999€, 21% zwischen 2000 und 2999€, 7% zwischen 3000 und 3999€ und 6% haben ein Einkommen über 4000€. Weitere 10% der Befragten machten keine Angabe zu ihrer Einkommenssituation.

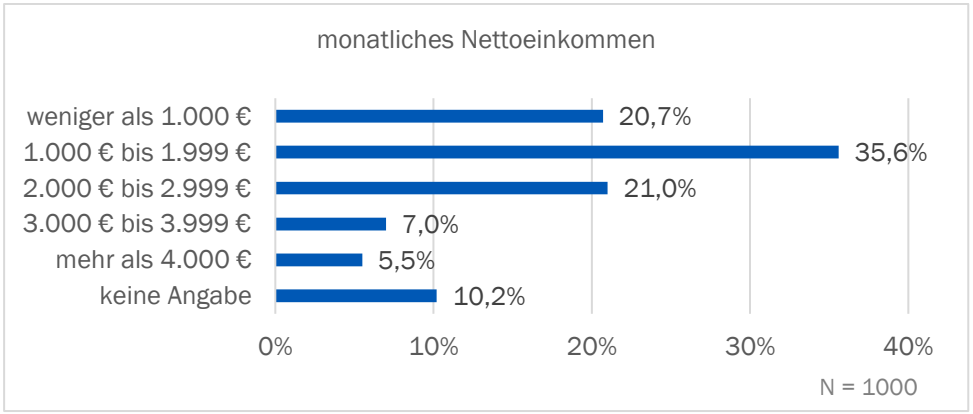


Abbildung 94: Monatliches Nettoeinkommen der Befragten

9.2.3.4 Berufliche Situation

42% der Befragten sind voll erwerbstätig, während 13% in Teilzeit und weitere 2% geringfügig oder unregelmäßig beschäftigt sind. 13% der Umfrageteilnehmer sind noch in Ausbildung, das heißt entweder Schüler/in, Student/in oder in betrieblicher Ausbildung oder Lehre. Weitere 19% geben an Rentner/in oder Pensionär/in zu sein. 9% des Samples sind nicht erwerbstätig.

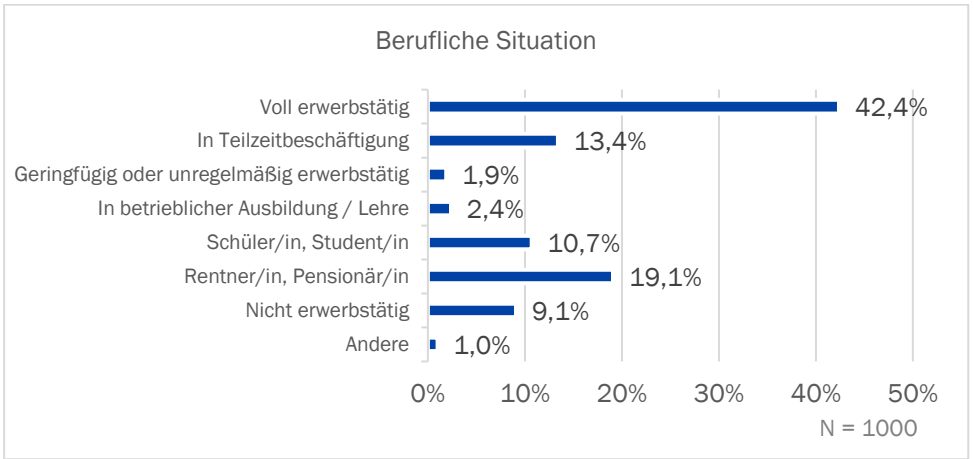


Abbildung 95: Berufliche Situation der Befragten

9.2.3.5 Höchster Bildungsabschluss

Hinsichtlich des höchsten Bildungsabschlusses ist das Sample durchmischt. 27% der Umfrageteilnehmer verfügen über einen Hauptschulabschluss und 30% über einen Realschulabschluss. 21% geben an eine Fachhochschulreife oder Abitur erlangt zu haben. Weitere 19% verfügen über ein abgeschlossenes Studium (Bachelor, Master oder Diplom). Ein geringer Anteil von 2% hat eine Promotion oder Habilitation abgeschlossen. 1% der Teilnehmer können keiner vordefinierten Bildungskategorie zugeornet werden und weitere 1% haben keinen Bildungsabschluss erlangt.

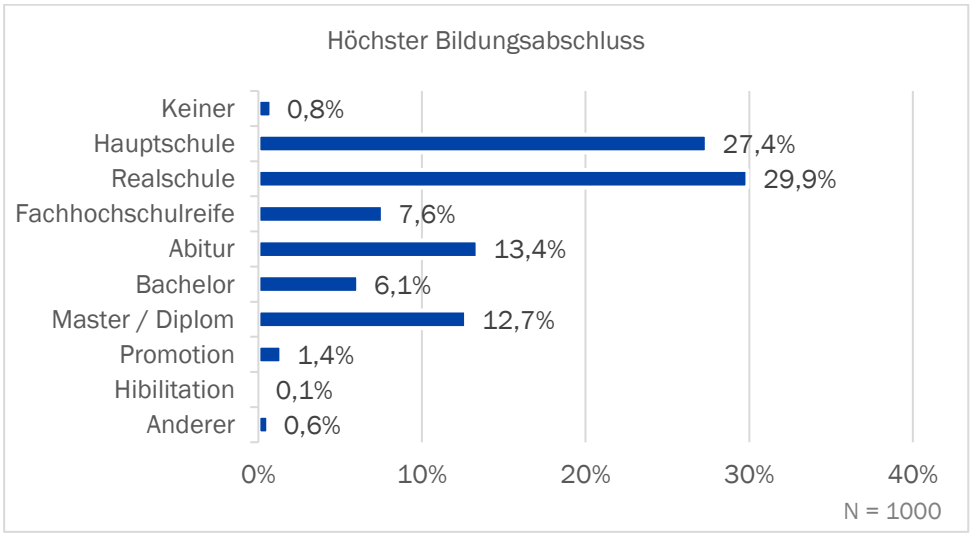


Abbildung 96: Höchster Bildungsabschluss der Befragten

9.2.3.6 Staatsangehörigkeit

96% der Befragten haben die deutsche Staatsangehörigkeit, während 4% eine andere Staatsangehörigkeit besitzen.

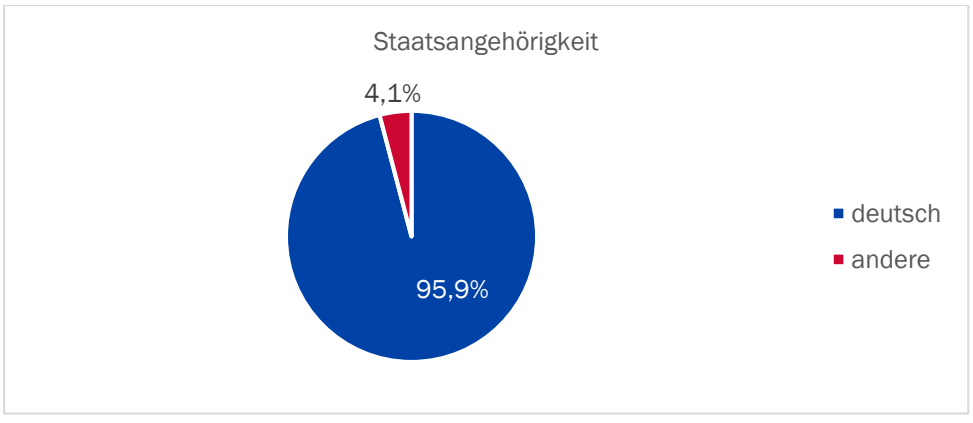


Abbildung 97: Staatsangehörigkeit der Befragten

9.2.3.7 Sprache

96% der Befragten geben an, dass sie in ihrem Haushalt hauptsächlich deutsch sprechen. Die andern 4% sprechen hauptsächlich eine andere Sprache.

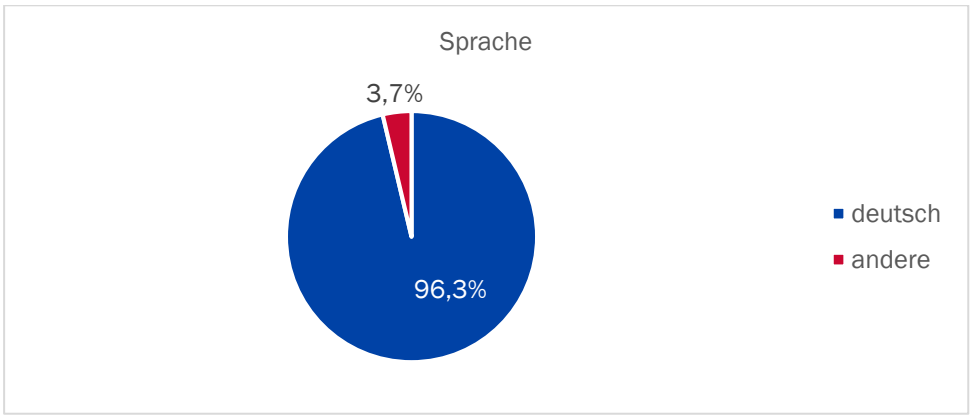


Abbildung 98: Sprache im Haushalt der Befragten

9.3 Befragung 2

9.3.1 Fragebogen

Parallel zur ersten Befragung wurden den Teilnehmerinnen und Teilnehmern zu Beginn mehrere Filterfragen gestellt. Hierzu zählen die Smartphone-Nutzung, Alter, Geschlecht und Wohnort. Wurde die Filterfrage zur Smartphonenutzung mit nein beantwortet, wurde der Fragebogen beendet. Die Fragen zu Alter, Geschlecht und Wohnort wurden zur Erreichung der Quoten parallel zur ersten Befragung verwendet. Im Folgenden werden die Einleitung sowie exakten Fragen zur Untersuchung der Themencluster der Informationstexte aufgelistet. Hierbei ist zu beachten, dass ein weiterer Teil der Befragung sich mit dem Thema der digitalen Verantwortung beschäftigte. Dieser Teil wird im Abschlussbericht nicht dargestellt.

Einleitung

Sehr geehrte Teilnehmerin, sehr geehrter Teilnehmer,

vielen Dank dafür, dass Sie an unserer Umfrage mitwirken!

Diese Umfrage wurde von einer Forschungsgruppe an der Quadriga Hochschule Berlin gestaltet.

Die Umfrage besteht aus zwei Teilen. Im ersten Teil geht es um digitale Verantwortung und im zweiten Teil um das Themenfeld Datenschutz und -sicherheit bei Webseiten und Smartphone-Apps.

Für die Umfrage sollten Sie sich ca. 15 Minuten Zeit nehmen. In der Befragung geht es um Ihre persönliche Meinung.

Die Umfrage wird ausschließlich von den Mitarbeiterinnen und Mitarbeitern des Forschungsprojekts ausgewertet.

Ihre Antworten sind hierbei vollkommen anonym und es sind keine Rückschlüsse auf Ihre Person möglich.

Teil 2: Datenschutz und Datensicherheit

Im zweiten Teil dieser Umfrage möchten wir genauer auf das Thema **Datenschutz und Datensicherheit** bei Webseiten und Apps eingehen.

Hierbei bitten wir Sie zuerst um Ihre Einschätzungen zu acht verschiedenen Themenfeldern. Bei jedem Themenfeld werden Sie zuerst gefragt, wie einfach es für Sie ist, Informationen zu diesem Thema zu finden. Danach werden Sie gefragt, wie wichtig es Ihnen ist, Zugang zu solchen Informationen zu haben.

1. **Datenschutzerklärungen**

Wie einfach oder schwierig ist es für Sie zu erkennen, ob die Datenschutzerklärungen von Webseiten oder Apps geltendes Recht einhalten?

	Sehr schwierig			Sehr einfach

Webseiten	0	0	0	0
Apps	0	0	0	0

Wie wichtig oder unwichtig ist es für Sie zu erkennen, ob die Datenschutzerklärungen von Webseiten oder Apps geltendes Recht einhalten?

	Sehr unwichtig			Sehr wichtig
Webseiten	0	0	0	0
Apps	0	0	0	0

Identifikation, das heißt die Möglichkeit Sie oder Ihr Gerät durch den Einsatz digitaler Technologien bei der wiederholten Nutzung von Webseiten oder Apps wiederzuerkennen

Wie einfach oder schwierig ist es für Sie zu erkennen, ob Webseiten oder Apps Sie bzw. Ihr Gerät (PC, Laptop oder Smartphone) anhand von Daten, die über Sie erhoben wurden, identifizieren?

	Sehr schwierig			Sehr einfach
Webseiten	0	0	0	0
Apps	0	0	0	0

Wie wichtig oder unwichtig ist es für Sie zu erkennen, ob Webseiten oder Apps Sie bzw. Ihr Gerät (PC, Laptop oder Smartphone) anhand von Daten, die über Sie erhoben wurden, identifizieren?

	Sehr unwichtig			Sehr wichtig
Webseiten	0	0	0	0
Apps	0	0	0	0

Zugriff

Wie einfach oder schwierig ist es für Sie zu erkennen, ob Webseiten oder Apps Zugriff auf Ihr Adressbuch oder Ihren Standort haben?

	Sehr schwierig			Sehr einfach
Webseiten	0	0	0	0
Apps	0	0	0	0

Wie wichtig oder unwichtig ist es für Sie zu erkennen, ob Webseiten oder Apps Zugriff auf Ihr Adressbuch oder Ihren Standort haben?

	Sehr unwichtig			Sehr wichtig
Webseiten	0	0	0	0
Apps	0	0	0	0

Sicherheit

Wie einfach oder schwierig ist es für Sie zu erkennen, ob Webseiten oder Apps sicher sind bspw. indem sie Verschlüsselung verwenden?

	Sehr schwierig			Sehr einfach
Webseiten	0	0	0	0
Apps	0	0	0	0

Wie wichtig oder unwichtig ist es für Sie zu erkennen, ob Webseiten oder Apps sicher sind bspw. indem sie Verschlüsselung verwenden?

	Sehr unwichtig			Sehr wichtig
Webseiten	0	0	0	0
Apps	0	0	0	0

Profilbildung

Wie einfach oder schwierig ist es für Sie zu erkennen, ob Webseiten oder Apps anhand Ihrer Daten ein Profil von Ihnen erstellen?

	Sehr schwierig			Sehr einfach
Webseiten	0	0	0	0
Apps	0	0	0	0

Wie wichtig oder unwichtig ist es für Sie zu wissen, ob Webseiten oder Apps anhand Ihrer Daten ein Profil von Ihnen erstellen?

	Sehr unwichtig			Sehr wichtig
Webseiten	0	0	0	0
Apps	0	0	0	0

Werbung

Wie einfach oder schwierig ist es für Sie, Werbung oder andere gesponserte Inhalte bei der Nutzung von Webseiten oder Apps zu erkennen?

	Sehr schwierig			Sehr einfach
Webseiten	0	0	0	0
Apps	0	0	0	0

Wie wichtig oder unwichtig ist es für Sie, Werbung oder andere gesponserte Inhalte bei der Nutzung von Webseiten oder Apps zu erkennen?

	Sehr unwichtig			Sehr wichtig
Webseiten	0	0	0	0
Apps	0	0	0	0

Personalisierte Werbung, das heißt Werbung, die auf Ihre Person oder Ihr Gerät zugeschnitten wird

Wie einfach oder schwierig ist es für Sie zu erkennen, welche der Daten, die über Sie auf Webseiten oder in Apps erhoben wurden, für personalisierte Werbung verwendet werden?

	Sehr schwierig			Sehr einfach
Webseiten	0	0	0	0
Apps	0	0	0	0

Wie wichtig oder unwichtig ist es für Sie zu erkennen, welche der Daten, die über Sie auf Webseiten oder in Apps erhoben wurden, für personalisierte Werbung verwendet werden?

	Sehr unwichtig			Sehr wichtig
Webseiten	0	0	0	0
Apps	0	0	0	0

Übertragung an Dritte

Wie einfach oder schwierig ist es für Sie zu erkennen, welche Ihrer Daten Webseiten oder Apps an Dritte weitergeben?

	Sehr schwierig			Sehr einfach
Webseiten	0	0	0	0
Apps	0	0	0	0

Wie wichtig oder unwichtig ist es für Sie zu erkennen, welche Ihrer Daten Webseiten oder Apps an Dritte weitergeben?

	Sehr unwichtig			Sehr wichtig
Webseiten	0	0	0	0
Apps	0	0	0	0

Abschlussfragen

Abschließend haben wir noch ein paar Fragen zu Ihrer Person.

Wenn man alle Einkünfte zusammennimmt: Wie hoch ist Ihr monatliches Einkommen?

Bitte geben Sie den monatlichen Netto-Betrag an, also nach Abzug von Steuern und Sozialabgaben. Rechnen Sie bitte regelmäßige Zahlungen wie Renten, Wohngeld, Kindergeld, BAföG, Unterhaltszahlungen usw. dazu.

- Weniger als 1000 €
- 1000 bis 1999 €
- 2000 bis 2999 €
- 3000 bis 3999 €
- Mehr als 4000 €
- keine Angabe

Wie ist Ihre derzeitige berufliche Situation?

- Voll erwerbstätig
- In Teilzeitbeschäftigung
- In betrieblicher Ausbildung / Lehre
- Schüler/in, Student/in
- Geringfügig oder unregelmäßig erwerbstätig
- Rentner/in, Pensionär/in
- Nicht erwerbstätig
- Andere: _____

Was ist Ihr höchster Bildungsabschluss?

- Hauptschulabschluss
- Mittlerer Schulabschluss (zum Beispiel Realschulabschluss)
- Fachhochschulreife (Abschluss einer Fachoberschule)
- Abitur (Hochschulreife)
- Bachelor
- Master / Diplom
- Promotion
- Habilitation
- Andere: _____

Vielen Dank für Ihre Teilnahme!

9.3.2 Beschreibung der Stichprobe

Im Rahmen der zweiten Online-Umfrage wurden insgesamt 1.000 Menschen mit Smartphone befragt. Ungewichtet weist die Stichprobe folgende Merkmale auf:

9.3.2.1 Alter

Die Altersquoten der ersten Befragung wurden als Grundlage für die Altersquoten der zweiten Befragung verwendet und bei der Stichprobenziehung durch das Marktforschungsinstitut zu Grunde gelegt. Deshalb stimmt die Altersverteilung der zweiten Befragung mit der der ersten Befragung überein. Der jüngsten Alterskohorte (14-19 Jahre) gehören 9% der Befragten an. 17% der Umfrageteilnehmer sind zwischen 20 und 29 Jahren, 16% zwischen 30 und 39 Jahren, 20% zwischen 40 und 49 Jahren, 17% zwischen 50 und 59 Jahren, 10% zwischen 60 und 69 Jahren und 12% sind 70 Jahre oder älter.

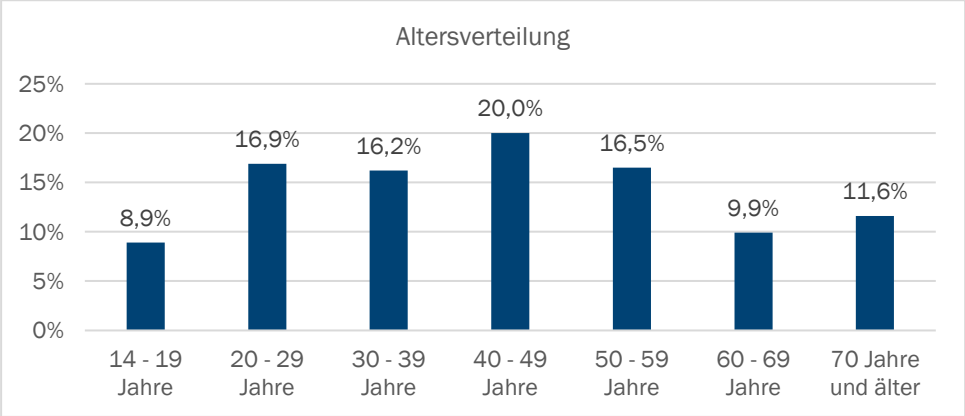


Abbildung 99: Alter der Befragten – Befragung 2

9.3.2.2 Geschlecht

Auch bei der Geschlechterverteilung wurde für die zweite Befragung die Verteilung der ersten Befragung zu Grunde gelegt. In der Stichprobe finden sich insgesamt 49,5% männliche Teilnehmer und 50,5% weibliche Teilnehmerinnen.

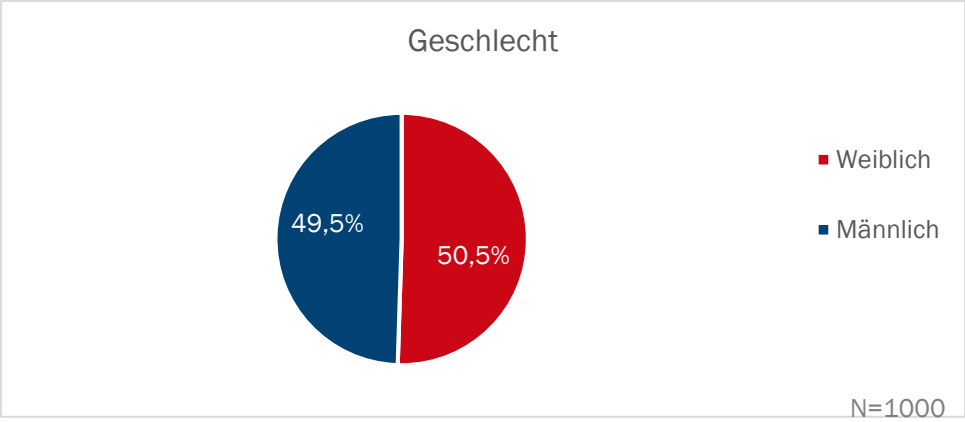


Abbildung 100: Geschlecht der Befragten – Befragung 2

9.3.2.3 Nettoeinkommen

Für das Nettoeinkommen ergab sich für die zweite Befragung folgende Verteilung. 12% der Teilnehmerinnen und Teilnehmer gaben an, dass sie weniger al 1.000€ zur Verfügung haben. 26% haben ein monatliches Nettoeinkommen zwischen 1.000 und 1.999€, 25% zwischen 2.000 und 2.999€ und 16% zwischen 3.000 und 3.999€. Zur Einkommenskategorie „mehr als 4.000€“ können weitere 20% der Befragten gezählt werden. Außerdem machen 11% der Teilnehmerinnen und Teilnehmer keine Angebe zu ihrem Einkommen.

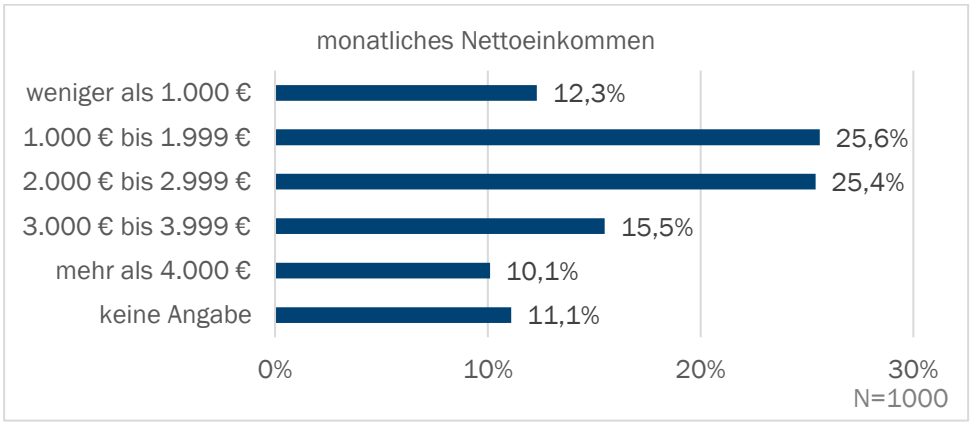


Abbildung 101: Monatliches Nettoeinkommen der Befragten – Befragung 2

9.3.2.4 Berufliche Situation

47% der Befragten sind voll erwerbstätig, während 13% in Teilzeit und weitere 1% geringfügig oder unregelmäßig beschäftigt sind. 13% der Umfrageteilnehmer sind noch in Ausbildung, das heißt entweder Schüler/in, Student/in oder in betrieblicher Ausbildung oder Lehre. Weitere 17% geben an Rentner/in oder Pensionär/in zu sein. 9% des Samples sind nicht erwerbstätig.

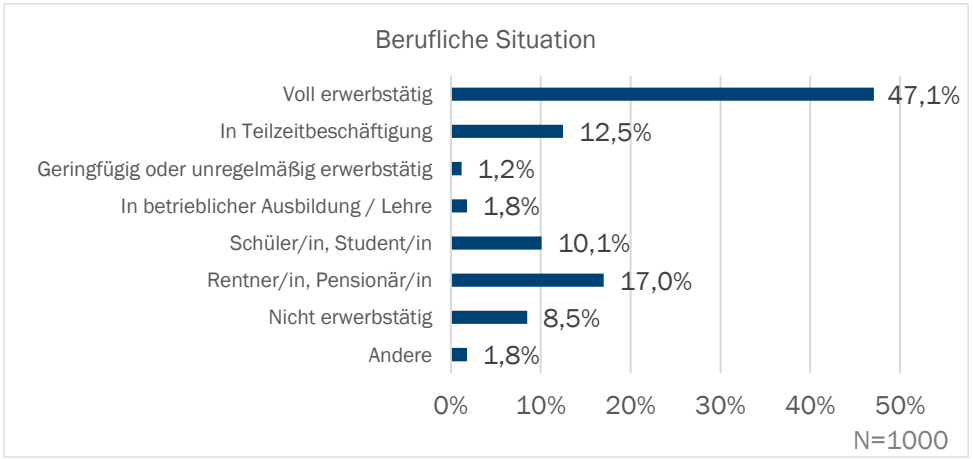


Abbildung 102: Berufliche Situation der Befragten – Befragung 2

9.3.2.5 Höchster Bildungsabschluss

12% der Umfrageteilnehmer verfügen über einen Hauptschulabschluss und 32% über einen Realschulabschluss. 30% geben an eine Fachhochschulreife oder Abitur erlangt zu haben. Weitere 24% verfügen über ein abgeschlossenes Studium (Bachelor, Master oder Diplom). Ein geringer Anteil von 1% hat eine Promotion oder Habilitation abgeschlossen. 2% der Teilnehmer können keiner vordefinierten Bildungskategorie zugeordnet werden. Hierzu zählen unter 1% die keinen Bildungsabschluss keinen erlangt.

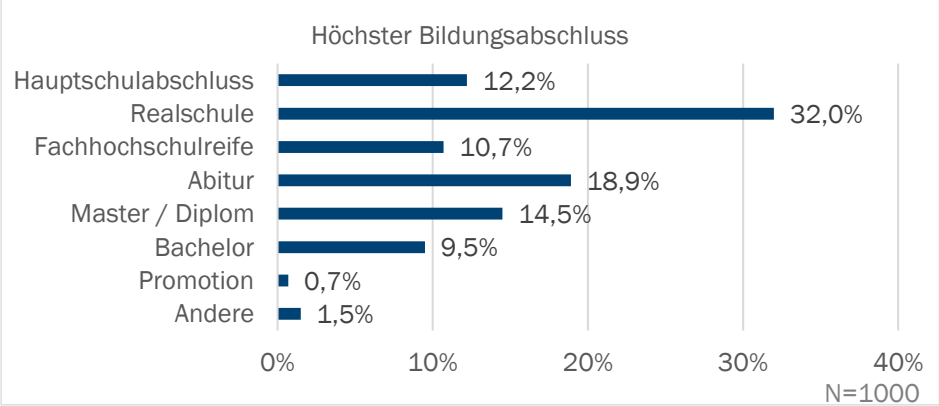


Abbildung 103: Höchster Bildungsabschluss der Befragten – Befragung 2

Die Staatsangehörigkeit und Sprache der Teilnehmerinnen und Teilnehmer wurde in der zweiten Befragung nicht abgefragt.

9.4 Fokusgruppe - Beschreibung der Stichprobe

	Gruppe 1	**Gruppe 2**	**Gruppe 3**
Geschlecht			
Weiblich in %	50%	50%	40%
Alter			
bis 35 Jahre	20%	30%	20%
Zwischen 35 und 60 Jahren	50%	40%	50%
Über 60 Jahre	30%	30%	30%
Anzahl Apps auf eigenem Smartphone	Median: 22,5 Mittelwert: 54,7	Median: 17,5 Mittelwert: 26,2	Median: 40 Mittelwert: 36
Kenntnisstand der Teilnehmerinnen und Teilnehmer beim Thema Datenschutz	Mittelmäßig, vereinzelte Experten	Mittelmäßig, vereinzelte Experten	Mittelmäßig, vereinzelte Experten
Bisherige Datenschutzmaßnahmen der Teilnehmerinnen und Teilnehmer	Virenschutz, datensparsames Verhalten jedoch auch unbekümmerte TN	Datensparsames Verhalten, Unterstützung durch Expertinnen und Experten in der Familie	Passwortschutz, datensparsames Verhalten

Tabelle 10: Beschreibung der Stichprobe in den Fokusgruppengesprächen

9.5 Experten Beta-Phase – Tutorial

Herzlich Willkommen beim DATENSCHUTZscanner!

Der DATENSCHUTZscanner möchte mit seinem Ansatz den Selbstdatenschutz im Smartphone wiederherstellen und Nutzerinnen und Nutzern mehr Transparenz und Kontrolle im Umgang mit ihren Daten ermöglichen.
Wir freuen uns über Ihr Feedback und möchten Ihnen mit diesem Tutorial eine Einführung in den Prototypen geben.
Im Tutorial zeigen wir Ihnen

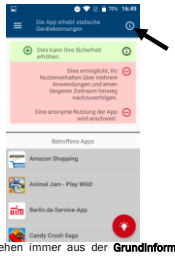
1. Wie man mit dem DATENSCHUTZscanner mehr über die Datenverarbeitung einer App erfahren kann.
2. Wie man mit dem DATENSCHUTZscanner erfährt, welche Datenverarbeitungen auf installierte Apps zutrifft.

Da wir uns derzeit in der Entwicklungsphase befinden, ist es möglich, dass Inhalte noch fehlerhaft sind und auch das Layout und die App-Funktionen Lücken enthalten. Wenn Ihnen Fehler auffallen oder Sie Verbesserungsvorschläge haben, teilen Sie uns diese bitte mit.

1 Startbildschirm

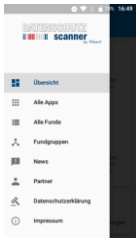
Auf dem Startbildschirm finden Sie erste Informationen und im unteren Bereich werden Ihnen die wichtigsten „Funde" angezeigt.
Funde sind der Kernbestandteil des DATENSCHUTZscanners. Sie sind relevante Datenverarbeitungen und Verstöße, die sich aus unserer technischen Analyse und der Prüfung der Datenschutzerklärung ergeben haben.

2 Funde



Funde bestehen immer aus der **Grundinformation**. Zusätzlich werden über das **Icon (i) Detail-Informationen** angezeigt, die das Verständnis unterstützen. Da der DATENSCHUTZscanner nicht pauschal bewertet, zeigt er sowohl **negative**, als auch **positive Konsequenzen** an. Am Ende entscheiden die Nutzerinnen und Nutzer durch Bewertung selbst, wie sie Verarbeitungen beurteilen und ob sie handeln.

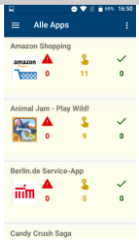
3 Menü



Die App wird über das Menü gesteuert. Unter „**Alle Apps**" finden Sie eine Liste der installierten Apps. Im Bereich „**Alle Funde**" erhalten Sie eine Liste aller Funde und in den „**Fundgruppen**" werden die einzelnen Funde nach Verarbeitungsarten sortiert. „**News**" sind bisher nur Platzhalter. Im Bereich „**Partner**" und „**Datenschutzerklärung**" finden Sie Details zum Team und unseren Datenverarbeitungen.

Tutorial 1: Wie man mit dem DATENSCHUTZscanner mehr über die Datenverarbeitungen einer App erfährt – Beispiel „Berlin.de Service-App"

T1.1

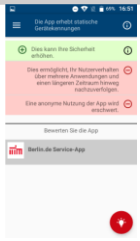


Öffnen Sie das **Menü** und wählen Sie „**Alle Apps**". Sie sehen eine Liste aller installierten Apps und in **gelb die Anzahl der Funde**. Die „Berlin.de Service-App" hat beispielweise 5 Funde. Durch Ihre persönlichen Bewertungen können Sie die Funde später negativ oder positiv bewerten. Diese Bewertung wird dann über die **roten** und **grünen** Zahlen angezeigt.
Klicken Sie nun bitte auf die „Berlin.de Service-App".

T1.2

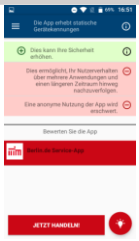
Die „**App-Details**" der „Berlin.de Service-App" zeigen alle Funde an, die unsere Analyse ergeben hat.
Beispielsweise wird hier angezeigt, dass die App **statische Gerätekennungen** erhebt.
Wenn Sie mehr Informationen zu diesem Fund erhalten möchten, klicken Sie diesen nun an.

T1.3



Nun erhalten Sie zusätzliche Informationen zu diesem Fund und können sich ein genaueres Bild über die möglichen Konsequenzen der Datenverarbeitung machen. Über das **Icon (i)** erhalten Sie weitere **Detail-Informationen**.
Im unteren Bereich können Sie die App nun bewerten.

T1.4



Finden Sie den Fund „statische Gerätekennungen" für die „Berlin.de"-App nicht in Ordnung, dann wischen das App-Feld von links nach rechts. Das **graue Feld** färbt sich **rot** und entspricht einer negativen Bewertung. Finden Sie die Erhebung von „statischen Gerätekennungen" in Ordnung, und möchten eine positive Bewertung vornehmen, so wischen Sie von rechts nach links. Das **graue Feld** färbt sich **grün**.

Abbildung 104: Experten Beta-Phase - Tutorial[133]

[133] Die beispielhaft verwendeten Namen und App-Logos dienen ausschließlich der Illustration; Aussagen über tatsächliche Datenverarbeitungen werden hierdurch nicht getroffen, ebenso wie sich entsprechende Rückschlüsse ausdrücklich verbieten.

10 Abbildungsverzeichnis

11 Tabellenverzeichnis