The Journal of Systems & Software 195 (2023) 111545

Contents lists available at ScienceDirect

The Journal of Systems & Software

journal homepage: www.elsevier.com/locate/jss

Privacy explanations – A means to end-user trust[☆]

Wasja Brunotte[a,b,*], Alexander Specht[a], Larissa Chazette[a], Kurt Schneider[a,b]

[a] Leibniz University Hannover, Software Engineering Group, Hannover, Germany
[b] Leibniz University Hannover, Cluster of Excellence PhoenixD, Hannover, Germany

ARTICLE INFO

Article history:
Received 7 April 2022
Received in revised form 26 July 2022
Accepted 12 October 2022
Available online 28 October 2022

MSC:
68-02
99-00

Keywords:
Explainability
Privacy definition
Privacy explanations
Online privacy
Privacy awareness

ABSTRACT

Software systems are ubiquitous, and their use is ingrained in our everyday lives. They enable us to get in touch with people quickly and easily, support us in gathering information, and help us perform our daily tasks. In return, we provide these systems with a large amount of personal information, often unaware that this is jeopardizing our privacy. End users are typically unaware of what data is collected, for what purpose, who has access to it, and where and how it is stored. To address this issue, we looked into how explainability might help to tackle this problem. We created privacy explanations that aim to help to clarify to end users why and for what purposes specific data is required. We asked end users about privacy explanations in a survey and found that the majority of respondents (91.6 %) are generally interested in receiving privacy explanations. Our findings reveal that privacy explanations can be an important step towards increasing trust in software systems and can increase the privacy awareness of end users. These findings are a significant step in developing privacy-aware systems and incorporating usable privacy features into them, assisting users in protecting their privacy.

1. Introduction

Personal data has long become a sort of virtual currency (Rana and Weinman, 2015; Patil and Shyamasundar, 2019). In 2017, a headline in *The Economist* stated that "the world's most valuable resource is no longer oil, but data" (Parkins, 2017; Wieringa et al., 2021). This has allowed a thriving new industry to emerge, known as *data brokers*. This lucrative and fast-growing industry treats data as a commodity. We generate enormous amounts of data every second, leaving digital traces of our online selves behind (Klitou, 2014). All this data is usually stored, merged, and evaluated (Wieringa et al., 2021; Schneier, 2015; Dinev, 2014). For instance, when driving a car, data about the speed and the strength with which the driver steps on the pedals is collected. This data can be used for routine diagnosis or for accountability purposes in the event of an accident. A simple photo contains information such as timestamp, location (GPS coordinates), camera information, and settings that are often collected, processed, and stored with the actual image data (Schneier, 2015).

Unfortunately, government agencies and private companies that dispose of our data do not always use appropriate mechanisms to prevent accidental or intentional privacy violations (Bowman et al., 2015). The "power" hidden in data has led institutions around us to rightly conclude something rather obvious: that this data has enormous value.

1.1. The privacy dilemma

Every time someone uses a software system, they consciously or unconsciously make a *trade-off* between the benefits of using the system and the data they provide during use. The reason for this data exchange (and thus the disclosure of personal information) is often motivated by access to personalized content, "free" information (Tun-Min et al., 2016), discounts (Barnett White, 2004), and loyalty programs (Earp and Baumer, 2003), as well as other economic incentives (Hann et al., 2002). Yet, this information disclosure usually happens without explicit (informed) consent, although this data is related to and belongs to the end user (Janssen et al., 2020).

The majority of websites (over 60% in Europe) rely on (cookie) *consent notices* to get visitors' consent to their data practices. However, because the implementations of these consent notices have substantial usability flaws, it is frequently unclear to the end user what data is being collected, stored, and for what purpose (Utz et al., 2019; Soe et al., 2020).

Hence, a more responsible approach to personal data is needed, both from the regulatory side and from the companies themselves. According to Garcia-Rivadulla (2016), companies

[☆] Editor: Prof. Raffaela Mirandola.
[*] Corresponding author at: Leibniz University Hannover, Software Engineering Group, Welfengarten 1, 30167 Hannover, Germany.
E-mail addresses: wasja.brunotte@inf.uni-hannover.de (W. Brunotte), alexander.specht@inf.uni-hannover.de (A. Specht), larissa.chazette@inf.uni-hannover.de (L. Chazette), kurt.schneider@inf.uni-hannover.de (K. Schneider).

https://doi.org/10.1016/j.jss.2022.111545
0164-1212/© 2022 Elsevier Inc. All rights reserved.

should not perceive this as a threat. Rather, they should see it as an opportunity to innovate in the context of privacy and gain consumer trust. In a study, Cummings et al. (2021) discovered that if users are given more information about how their data will be used, they are more willing to provide this data.

1.2. Data economy and privacy awareness

With all these factors in mind, it is crucial to find alternative solutions to this problem. As Bowman et al. (2015) state, "engineers have a responsibility to the rest of society". Therefore, even if there is no "right" answer about one's right on the level of privacy or on how privacy could be most effectively protected, software engineers should actively tackle this challenge and find answers to the existing open questions surrounding this topic.

Following this line of thought, the "principle of minimum asymmetry" should be considered (Jiang et al., 2002). According to this principle, "a privacy-aware system should minimize the asymmetry of information between data owners,[1] data collectors,[1] and data users".[1] This should be done by "*decreasing* the flow of information from data owners to data collectors and users" and "*increasing* the flow of information from data collectors and users back to data owners" (Jiang et al., 2002).

To address today's challenges regarding privacy (data economy and benefits), it is critical to develop systems that are *privacy-aware* in design, incorporate *usable* privacy features, and enter into a transparent dialogue with end users about data practices.

1.3. Privacy explanations as a solution

So far, privacy policies are the primary channel to inform users about data practices of a service provider. However, privacy policies are insufficient when it comes to informing users since they are too long, too vague, and the information required can often only be interpreted with legal background knowledge (Brunotte et al., 2022b; Jensen and Potts, 2004; McDonald and Cranor, 2008; Pollach, 2007; Reidenberg et al., 2015). Up to this moment, there is a lack of a user-centered solution to explain privacy-related aspects in an appropriate and understandable way.

Explainability is a non-functional requirement (NFR) that is increasingly seen as a means to mitigate a system's lack of transparency and provide an understanding of a system's behavior among end users by giving explanations and disclosing information (Chazette et al., 2021; Köhl et al., 2019). Thereupon, explanations might be a means to inform end users about a system's data practices. For instance, if an app needs access to a user's location, explanations can inform the user about the purpose and scope of the data collection.

To contribute to the research of privacy in software engineering, we follow our research agenda (Brunotte et al., 2021), by employing the concept of explainability to bridge the gap between the process of disclosing personal information and the lack of transparency involved in this process. To this end, we conducted an online survey to assess whether there is a need among end users to receive explanations of privacy aspects, how they perceive such explanations, and whether they have an influence on end users' trust toward a system. The results of our study show that respondents are interested in privacy explanations. They consider them beneficial, and state that privacy explanations might contribute to increasing trust in software.

This paper is structured as follows: in the following Section (2), we present background and related work. In Section 3, we present our research questions (RQs), and outline the chosen research design. In Section 4, we present the findings of our survey, in Section 5, we discuss our results and propose a forecast about future work (7). In Section 6 we discuss threats to validity. Finally, we conclude our paper in Section 8.

2. Background and related work

In this section, we define terms and provide background information that are necessary for the further understanding of this work. For this purpose, we will start discussing the different dimensions of *the concept of privacy* (2.1). Second, we delineate privacy and *online privacy* in 2.2. In doing this, we want to gain a better understanding of what these concepts imply in terms of quality aspects (NFRs) for software systems. In 2.3, we define the term privacy explanation. In 2.4, we define trust and trustworthiness. Although both terms are often used interchangeably, we would like to show why it makes sense from our point of view to distinguish the two terms from each other. We conclude this background section with related work in 2.5.

2.1. The concept of privacy

Privacy is a normative concept. It is not something new to our modern times but has existed for a very long time and is deeply rooted in sociological, philosophical, legal, political, and economic traditions (Nissim and Wood, 2018). In the past, when humans were still hunters and gatherers, it was of crucial importance to know where and when there was ripe fruit or where the next water source was located. Sharing this information only with certain individuals of one's own community could potentially ensure a group's survival (Harari, 2015).

Even though scholars from different disciplines have investigated the concept of privacy from different perspectives, there is still no unified view or definition regarding this concept (Renaud and Gálvez-Cruz, 2010; Moore, 2003; Yao, 2011; Krishna, 2020; Introna, 1997; Newell, 1995).

We provide an overview of authors' interpretations of privacy and its impact on individuals in Table 1. All of these views and expressions have one thing in common: that privacy is about the control of information about personal matters and the creation of private spaces, whether physical or mental. To summarize, privacy enables us to enter a state in which we can withdraw from society, either physically, mentally, or both. This withdrawal is socially tolerated, important for our well-being, and essential for a healthy society.

We combine the above-mentioned concepts of the various authors and scholars who have studied the notion of privacy in a definition. We want to achieve two goals with this definition: *(a)* compile a concise working definition that captures the many nuances of privacy and *(b)* build a shared understanding of what privacy actually means.

Definition 2.1: Privacy

Privacy is a right, a claim, and a state in which an individual independently sets their boundaries and determines what *personal matters*[a] they wish to share with or withhold from other individuals or society. These boundaries can be created through physical or psychological means. During this voluntary and temporary withdrawal, the individual is protected within these boundaries from unwanted intrusion (physically or mentally), embarrassment, judgment, discrimination, accountability, and societal norms and constraints.

[a] Personal matters refer to personal information, thoughts, feelings, and habits.

[1] Jiang et al. (2002) defined the terms data owner, data collector, and data user in their work, which we adopt here. *Data owners* are the individuals whose data is being used or accessed (e.g., end users). *Data collectors* are individuals or systems that collect information about data owners. *Data users* are individuals or systems that use (process) this information.

Table 1
Overview of authors' interpretations of privacy and its impact on individuals.

Authors' interpretations of privacy and its impact	Reference
Privacy is "a legal right" and "the right to be let alone"	Warren and Brandeis (1890)
"Privacy is structured by the answer . . . to the questions 'who are the persons you wish to exclude from having this knowledge?'"	Bates (1964)
"Privacy is the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others"	Westin (2015)
"The desire by each of us for physical space where we can be free of interruption, intrusion, embarrassment, or accountability and the attempt to control the time and manner of disclosures of personal information about ourselves"	Smith (2000)
"The right to privacy exists because democracy must impose limits on the extent of control and direction that the state exercises over the day-to-day conduct of individual lives"	Rubenfeld (1989)
"A state in which persons may find themselves"	Velecky (1978)
Influence of a person's well-being	Jourard (1966), Allen (1988), Moore (2003), Klopfer and Rubenstein (1977)
It is important for our mental and physical health	Petronio (2002)
Privacy is important to grow personally and its autonomy leaves room in determining one's own path in life	Moore (2003)
Privacy is important for the endurance of a strong society	Jourard (1966)
It acts like some sort of glue that binds individuals together in healthy relationships	Bräunlich et al. (2021)
A loss of privacy is not only unsettling but also represents an insult to a person's dignity, independence, and integrity	Bhave et al. (2020), Bloustein (1964)

Notwithstanding that privacy should be considered as a right of every individual (Warren and Brandeis, 1890; McCloskey, 1980; De Terwangne, 2012), there must also be regulatory entities that monitor and defend the strict observance of this right by legal means.

2.2. *Online privacy*

Our management of personal information has required ongoing adaptation and revision due to the rapid advancement of information and communication technology. For instance, printing technology simplified the reproduction and distribution of private information. The lines between personal and public life are blurring and shifting as a result of digitization and Internet use. We communicate using e-mail, messengers, and social media; seek for answers to private and sensitive questions using search engines. As a result, we leave a substantial quantity of digital evidence about our routines, opinions, and attitudes behind, making our privacy heavily dependent on how these systems are designed (Brunotte et al., 2022b; Cavoukian et al., 2009).

In the offline world (the physical world), we are mostly in control of our own privacy and protected by the normative concept of privacy, backed by social norms and legal traditions. For instance, if two friends are chatting in the town square, at most, bystanders can overhear portions of what they are saying. However, they can choose to stand further apart to ensure that their conversation is private. However, in the online world (the internet/cyberspace), this strategy is inapplicable and ineffectual. The amount of communication data saved here is significantly greater than a brief chat in the town square since it is stored in redundant ways (Klitou, 2014). To gain control over privacy in cyberspace, users must *actively* take care of their privacy themselves. Users cannot rely on legal systems and cannot expect other users to comply with their social and cultural norms (Yao, 2011).

Effective self-protection involves being conscious about and taking measures toward one's own privacy protection, as well as implies a certain technical knowledge about what is at stake, what is relevant to protect, and how to do it. However, not everyone has the necessary knowledge to identify privacy issues, and actively protecting one's privacy is also thought to be difficult and time-consuming (Rudolph et al., 2018).

In our definition, *privacy* is an individual's ability to control their physical or mental presence and their right to mental or physical seclusion. In contrast, *online privacy* is about an individual's control over their personal information in virtual space and their right to withhold this information. To this end, we define online privacy as:

Definition 2.2: Online Privacy

Online Privacy is a right, a claim, and a state in which an individual (data owner) sets their boundaries and determines (decide or control) what *privacy aspect*[a] they wish to share with data collectors and data users, by whom it may be accessed, and at what point in time this occurs.

[a] A privacy aspect refer to personal information, thoughts, and feelings. With respect to online privacy but especially data or information about a person. Examples: name, address, bank data, GPS location, etc.

2.3. *Privacy explanations*

Concerning one's privacy in software, long privacy policies or short privacy notices are often the only available sources of information to end users where they can (possibly) find out what happens to their data. End users rarely read or understand privacy policies (Brunotte et al., 2022b; Jensen and Potts, 2004; Pollach, 2007). As a result, users rarely benefit from privacy policies in terms of learning about their privacy. Therefore, users need a different form of clarification and transparency with respect to their online privacy.

Explainability is seen as an appropriate solution to mitigate the lack of transparency of a system (Chazette et al., 2019; Jasanoff, 2017; Richardson and Rosenfeld, 2018), it has an impact on the relationship of trust in a system and may lead to more end user acceptance (Chazette et al., 2021). Privacy explanations can inform users what a system will or will not do with their personal data.

In this work, a privacy explanation does ***not*** mean a privacy policy, privacy statement, or a privacy notice in the usual sense. We adopt the definition of explainable systems by Chazette et al. (2021) to define a *privacy explanation* as follows:

Definition 2.3: Privacy Explanation

A **privacy explanation** is a corpus of information I that a system S gives to an addressee A in context C to explain the purpose P for using a privacy aspect X.

The explanation I is intended to provide an explanation to the end user (addressee A), i.e., a reason why the user's privacy-related information (privacy aspect X) is needed. This could be, for example, why a smartphone app needs the user's location. It is important that this explanation provides a rationale (the purpose P) for why X is needed and does so in a transparent and understandable way. This explanation might be expressed in text, graphics, audio, or any combination of those. The context is the situation in which an explanation is given, consisting "of the interaction between a person, a system, a task, and an environment" (Chazette et al., 2021).

2.4. Trust and trustworthiness

Trust also plays an essential role in requirements engineering (RE), and system design (ISO Central Secretary, 2016; Giorgini et al., 2004; Elahi and Yu, 2009). Many see the concept of explainability as an appropriate means to amplify trust in a system, respectively stakeholder trust (Chazette et al., 2021; Chazette and Schneider, 2020; Langer et al., 2021). Explainability has been identified as an NFR in the RE community (Chazette et al., 2021; Chazette and Schneider, 2020; Köhl et al., 2019), and it is often connected with trust in the literature due to its potential to increase trust (Nagulendra and Vassileva, 2016; Chakraborti et al., 2019; Dahl, 2018; Floridi et al., 2018). In light of this, it might be more appropriate to engineer and elicit requirements for explainability than to have requirements for trust directly.

According to Kästner et al. (2021) is *trust* "an attitude a stakeholder holds *towards* a system". In contrast, the authors describe *trustworthiness* as "a property of a system: intuitively, a system is trustworthy for a stakeholder when it is warranted for the stakeholder to put trust in the system". In light of this, a system should "work properly" in a given context. Especially with regard to privacy, it is important to consider and differentiate trust and trustworthiness because if end users are to trust systems, they must also be sure that these systems are trustworthy.

2.5. Related work

Houghton and Joinson (2010) conducted a survey in which participants were asked to provide information on privacy concerns regarding social networks. The participants' biggest concern was the loss of control over their data, followed by the risk that supposed friends could turn out to be fraudsters.

Anton et al. (2010) asked participants in surveys about their privacy concerns when using software systems. The first survey began in 2002 and another survey was conducted in 2008. The result was that participants' concerns did not change, but the level of concern increased.

Wilkowska et al. (2020) conducted a survey with users of a daily living app and an application for people suffering from dementia. In their study, the subjects were asked, among other things, who they would trust with their data and whether the storage of the data was of interest. The result was that doctors and family members were perceived as trustworthy. When it comes to storing data, information about where and for how long the data is stored was considered essential. The storage of data in the cloud was disagreed with by the majority of participants.

The study by Wirth et al. (2021) examines the extent to which individual laziness affects privacy. The results showed that individuals who are predominantly lazy are more likely to disclose their data as well as not change their privacy settings, despite the fact that changes were made to data practices. Lazy people are also more likely to share their data. According to the study, however, hardworking people are more likely to be responsible with their data.

In the domain of artificial intelligence (AI), explainability has long played an important role (eXplainable Artificial Intelligence, XAI).

In Elahi et al. (2021) and Smart (2021), approaches are being pursued to improve people's understanding of data protection. AI systems are used for this purpose. Elahi et al. (2021) focus lie on elderly persons to prevent a cognitive overload while using different ambient assisted living systems. Smart (2021) provide an overview of different strategies to avoid privacy threats.

The works (Dai et al., 2022) and Wu et al. (2022) deal with privacy, fairness and explainability in the context of graph neural networks (GNNs). The authors mention in Dai et al. (2022) a positive interaction between explainability and privacy and that there is also a connection with trust. However, Wu et al. (2022) refers to possible problems in terms of privacy, when providing explanations for whitebox GNNs, since the model parameters are accessible.

A huge amount of research in the domain of XAI considers aspects such as fairness, trust, and privacy (Barredo Arrieta et al., 2020; Sheth et al., 2021; Balkir et al., 2022; Amparore et al., 2021; Tjoa and Guan, 2021; Mehdiyev et al., 2021). Here, as in the studies mentioned above, the focus was less on the end user and how to offer them information about data practices in a meaningful way.

In 2019, a study on the topic of explainability was conducted by Chazette et al. (2019). Their work examined the influence of explanations in a navigation application. Chazette et al. come to the conclusion that explanations can significantly increase usability, but unnecessary information can have a negative effect.

Cummings et al. (2021) used a survey to investigate whether users are more willing to share data if the disclosure of certain information is protected by differential privacy techniques. Their results showed that descriptions related to data use had a positive impact on respondents' willingness to share their data.

Furthermore, a large body of research was done in terms of privacy policies and how to communicate them in a more comprehensible way to end users (Brunotte et al., 2022a; Earp et al., 2005; Jensen and Potts, 2004; McDonald et al., 2009; Khan et al., 2020; Keymanesh et al., 2021; Chang et al., 2019; Nokhbeh Zaeem et al., 2020).

With our study, we would like to contribute to closing a research gap regarding the information asymmetry between end users and software systems. Our goal is to investigate to what extent the concept of explainability can be employed to inform end users in a simple, comprehensible, and satisfying way how their personal data is used.

3. Research goal and design

Our research goal was to examine the perception of end users with respect to privacy explanations. In particular, we focused on the influence of such privacy explanations, whether they may foster end user trust toward a system and whether they may play a role in increasing end users' privacy awareness. To this end, we formulate the goal of our research according to the goal definition template by Wohlin et al. (2012).

> **Goal definition:** We *analyze* end users' perceptions about the need for privacy explanations in software systems *for the purpose of* investigating whether explanations might influence

> the level of trust *from the point of view of* end users *in the context of* an online questionnaire.

Based on our goal definition, we framed our study into the following research questions (RQs).

> **RQ1:** How concerned are end users about their privacy and what threats regarding privacy are they worried about?
>
> **RQ2:** How do end users perceive explanations with respect to privacy aspects?
>
> **RQ3:** How are explanations of privacy aspects related to the concept of trust in a software system?

Online privacy is a significant issue because so much of modern life takes place online. As a result, we should concentrate on the creation of end user-centered privacy-aware systems. Accordingly, **RQ1** focuses on concerns and worries of end users with respect to their privacy. With this question, we wanted to get a general picture of what privacy risks users currently feel exposed to. Furthermore, this data should provide insight into whether end users' issues, or some of their concerns, can be mitigated by privacy explanations.

RQ2 focuses on how end users perceive privacy explanations, e.g., when an app communicates to its users why certain sensor data (location, etc.) from a smartphone is needed. On the one hand, we wanted to know whether users see the need for explanations when an app or service requires some privacy-related information. On the other hand, we wanted to know how users perceive it when a system provides an explanation regarding the use of a privacy aspect. To address this, we provided the following hypothetical scenario to survey participants:

> **Hypothetical Situation:** You heard from an acquaintance about a new app that lets you plan sightseeing tours or day trips for cities around the world. You decide to download this app to your smartphone for your upcoming city trip. However, when you start the app for the first time, it asks whether your location can be used and also asks for your date of birth. You are not sure about the reason, as the app does not give you any further information about the usage of your data.

In a first step, we asked participants if they would generally be interested in an explanation of *why* the app asks for the data. In a second step, we gave the participants an exemplary explanation of the privacy aspects (regarding use of the location and date of birth). We then asked whether the explanations were perceived as useful. Finally, we asked the participants what information a privacy explanation should contain.

With **RQ3**, we want to find out how privacy explanations and end user trust are related. Therefore, we asked the participants what are the benefits of privacy explanations. These findings might enable us to understand the needs and expectations of privacy explanations and how to meet them. Furthermore, we wanted to investigate whether privacy explanations are a suitable means of providing transparency regarding data practices, since both understanding and transparency are quality aspects that can foster user trust in a system (Chazette et al., 2021; Köhl et al., 2019; Ehsan et al., 2019).

In order to answer the RQs, we structured our research design as shown in Fig. 1. Each phase will be described in the following sub sections.

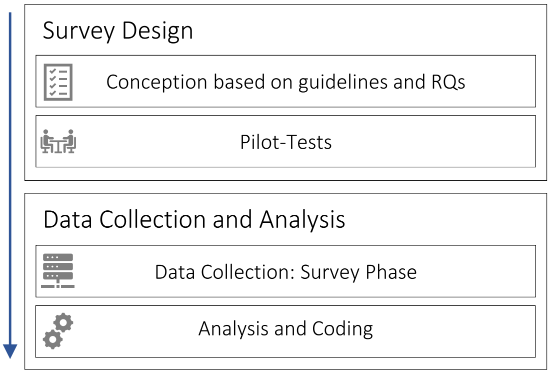

Survey Design
Conception based on guidelines and RQs
Pilot-Tests
Data Collection and Analysis
Data Collection: Survey Phase
Analysis and Coding

Fig. 1. Overview of the research design.

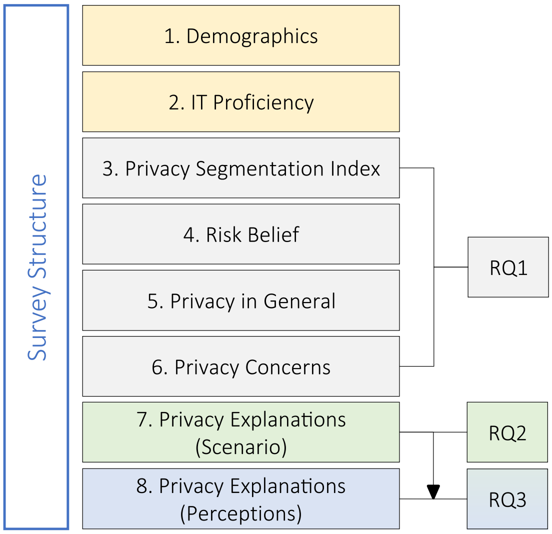


Fig. 2. Overview of the survey structure.

3.1. Survey design

To ensure the quality of our survey, we followed established guidelines for survey design (Sudman and Bradburn, 1982; Jacob et al., 2014; Groves et al., 2011). In order to avoid response bias, we took several actions in accordance with the guidelines. We tried to avoid leading questions, used precise and simple language, kept our questions short and clear as well as tried to use balanced and equal response categories. Furthermore, we informed the participants (in written form) that the survey is anonymous and they should answer honestly, as there are no right or wrong answers.

We defined the survey structure in line with our RQs. The survey started with a brief introduction and contained eight main parts with a total of 34 questions (30 multiple choice, four open-ended). Some of the multiple choice questions were also given the option for respondents to formulate their own answer text if none of the given answer options appealed to them. The structure of the survey, including the parts that answer the respective research questions, is depicted in Fig. 2.

The purpose of the first part (demographic questions) was to help to identify demographic factors that might have an impact on the respondents' answers. The second part contained questions to assess the respondents' experience with information technology (IT). The third part contained questions to collect respondents' Privacy Segmentation Index (PSI, Section 3.3.1). In the fourth part, we asked questions to help to assess participants' risk belief (Section 3.3.2). The fifth part comprised general questions about participants' privacy behavior. In the sixth part, participants were asked to describe situations in which they had concerns about their privacy when using software. In the seventh part of the survey, we gave the subjects a hypothetical scenario along with privacy explanations and asked questions regarding their perceptions. In the last part, we asked participants in general about possible benefits of privacy explanations as well as when a system should give such an explanation.

We conducted four rounds of pilot testing to assess the survey's quality. Two of these took place with members of our research group and two with candidates of the target population. Based on these pilots, we applied some minor corrections (e.g., addition of information regarding the interpretation of certain questions, minor text changes for better understanding).

3.2. Data collection and analysis

The data was collected via a web-based questionnaire. The survey was created with the survey tool LimeSurvey and hosted on our university's servers.

3.2.1. Data collection

Data collection took place over two months, starting in June 2021. We distributed the survey through many means, including academic mailing lists, Facebook, and Twitter, and we invited our personal network to share the survey with their networks. Our target group was adult end users with different occupations and IT knowledge since we wanted to understand the perception of end users with different backgrounds on this topic. Because of our sampling strategy (contact networks mostly concentrated in Germany, Brazil, and abroad), we provided our survey in three languages: English, German and Portuguese. We expected that a large part of the participants would come from Brazil and Germany.

3.2.2. Analysis and coding

We applied qualitative and quantitative analysis techniques to the survey results. We exported the results to spreadsheets in order to calculate descriptive statistics. For the open-ended questions, we applied a qualitative data analysis consisting of an open coding approach, as described by Saldaña (2013). According to Saldaña, coding is "one way of analyzing qualitative data" and it "transforms qualitative data into quantitative data, but it does not affect its subjectivity or objectivity" (Seaman, 1999).

We applied two consecutive coding cycles. First, we used *In Vivo Coding*, it is a method "to preserve participants' meanings of their views and actions in the coding itself" (Charmaz, 2014). In vivo codes serve as symbolic markers for the speech and meaning of the statements made by the respondents. We identified the essential passages of text in answers given by respondents in relation to the questions asked. A single response could result in more than one code. This was dependent on the length of the response as well as its meaning.

In the second coding cycle, we used *Pattern Coding*. In this type of coding, summaries are grouped into smaller sets, constructs, or themes (Miles and Huberman, 1994). For this purpose, categories were formed to reflect the meaning based on the codes. While forming these categories, we tried to preserve respondents' opinions and avoid over-interpretation. These categories can help to understand implicit meanings and actions. In addition, comparisons can also be made between the data and the categories resulting from the codes.

The coding procedure was conducted independently by the first two authors of this work. In cases of discrepancies, we discussed the differences until we reached consensus. We used Cohen's Kappa statistics (Cohen, 1968) to assess the reliability of the coding procedure. The resulting value of $\kappa = 0.87$ showed an almost perfect agreement (Landis and Koch, 1977).

3.3. Privacy segmentation index and risk beliefs

In our survey, we classified participants based on the Privacy Segmentation Index (PSI) as well as determined their Risk Beliefs. In the following, we describe the meaning of PSI and Risk Beliefs and how we determined them.

3.3.1. Privacy segmentation index

Westin developed the PSI to classify consumers according to their privacy concerns (Kumaraguru and Cranor, 2005). Although the PSI is related to a consumer perspective, it is adopted in broader contexts (Woodruff et al., 2014; Consolvo et al., 2005; Tun-Min et al., 2016). With this in mind, we have also collected the PSI to capture respondents attitude towards their privacy concerns. The following statements are included in the PSI where respondents express their level of agreement through a 7-point Likert scale (*strongly disagree* to *strongly agree*):

P1. Consumers have lost all control over how personal information is collected and used by companies.
P2. Most businesses handle the personal information they collect about consumers in a proper and confidential way.
P3. Existing laws and organizational practices provide a reasonable level of protection for consumer privacy today.

Based on the responses, participants can then be classified one of three categories: *Privacy Fundamentalists*, *Privacy Pragmatists*, and *Privacy Unconcerned*. The representative descriptions of these categories are given in the 2002 Harris report (Westin, 2002) as follows:

Privacy Fundamentalists: This group sees privacy as an especially high value, rejects the claims of many organizations to need or be entitled to get personal information for their business or governmental programs, thinks more individuals should simply refuse to give out information they are asked for, and favors enactment of strong federal and state laws to secure privacy rights and control organizational discretion.

Privacy Pragmatists: This group weighs the value to them and society of various business or government programs calling for personal information, examines the relevance and social propriety of the information sought, wants to know the potential risks to the privacy or security of their information, looks to see whether fair information practices are being widely enough observed, and then decides whether they will agree or disagree with specific information activities—with their trust in the particular industry or company involved a critical decisional factor. The Pragmatists favor voluntary standards and consumer choice over legislation and government enforcement. But they will back legislation when they think not enough is being done – or meaningfully done – by voluntary means.

Privacy Unconcerned: This group does not know what the "privacy fuss" is all about, supports the benefits of most organizational programs over warnings about privacy abuse, has little problem with supplying their personal information to government authorities or businesses, and sees no need for creating another government bureaucracy (a "Federal Big Brother") to protect someone's privacy.

According to Westin, the classification is as follows: Privacy Fundamentalists agree with P1 and disagree with both P2 and P3. Participants who disagree with P1 and agree with P2 and P3 belong to the Privacy Unconcerned class. The remaining participants can be classified as Privacy Pragmatists.

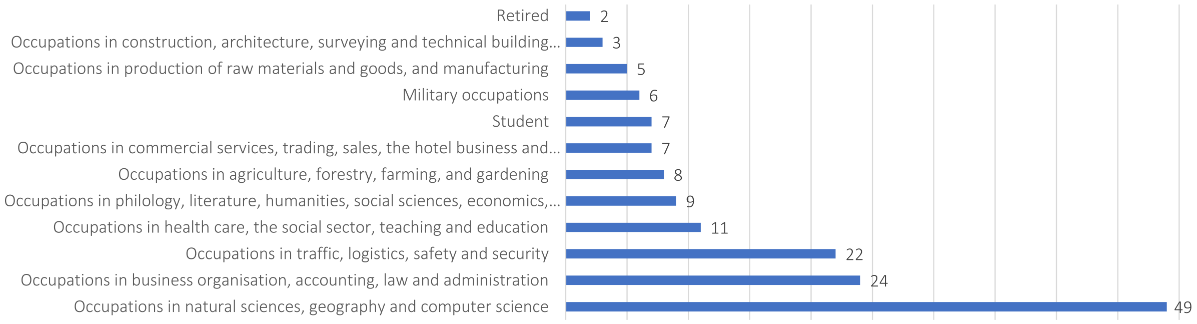


Fig. 3. Overview of the occupations.

3.3.2. Risk beliefs

Risk belief is a metric that can be used to quantify the level of risk a person perceives by sharing their information online. Tsai et al. (2006) used this metric in their work. We adopted their approach to calculate a **risk score** for each subject in the survey. Our aim has been to use the risk belief to possibly make a better assessment of the individual participants. For this purpose, we modified the questions of the risk beliefs a bit and formulated the questions not only in relation to online shopping, but in relation to online services in general. We asked respondents a total of four closed-ended questions for this purpose.

Q1. I feel safe giving my personal information to online services (such as online stores) and/or apps.*
Q2. Providing online services or apps with personal information causes too many concerns.
Q3. I generally trust online companies with handling my personal information, e.g., my purchase history.*
Q4. How concerned are you about threats to your personal privacy online today?

For questions one through three, we used a 7-point Likert scale to determine the level of agreement (*strongly disagree* to *strongly agree*). For the fourth question, we measured the level of concern with a 5-point Likert scale (*not at all concerned* to *extremely concerned*). Responses were scored according to the scales, with scores inverted for questions one and three to reflect feelings of concern (the higher the score, the higher the perceived risk). In order to map the 5-point Likert scale to the 7-point Likert scale, we weighted the items by 1.5. We see this mapping as justified and the participants' statements as not distorted. A chronbach's α (Cronbach, 1951) value of 0.76 confirmed the reliability of the 7-item scale (George and Mallery, 2009).

4. Results

In this section, we report the results related to our three RQs. In the first subsection, we start by presenting the participants' demographics. The subsequent subsections are devoted to each RQ, presenting the related results, and answering each RQ.

4.1. Demographics

We received a total of 209 responses. From the 209 participants who responded our survey, 155 completed the survey. For our data analysis, we considered only the 155 valid (complete) responses. Most respondents come from Germany (67.1%) and Brazil (21.9%). Ages ranged from 19 to 92 (M = 39, SD = 14.1).

61.9% of the participants identified themselves as male and 37.5% as female, and one (0.6%) to another gender. The majority of the participants work in the field of "natural science, geography and computer science" (32.03%) as shown in Fig. 3. Two respondents did not answer this question.

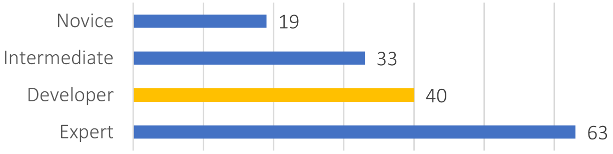


Fig. 4. IT skill levels of the respondents.

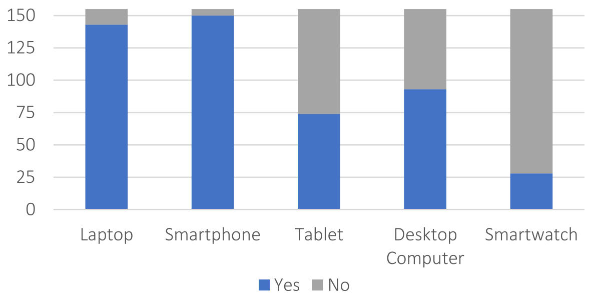


Fig. 5. Usage of different devices.

▸ *IT proficiency.* In order to assess respondents' IT proficiency, we included self-assessment questions where the respondents had to indicate whether they are able to perform certain tasks and whether they are familiar with certain IT terms.

As shown in Fig. 4, the majority of the survey respondents (66.5%, developers and IT experts) claimed to have a high level of IT proficiency. We have distinguished between *developers* and *experts* because developers not only have profound IT knowledge but they also have sound knowledge about the internals and other programming aspects of software systems.

The results show that, in general, respondents are comfortable working with software systems. We grouped respondents into the group *intermediate* if they claim to have at least significant software skills (such as creating functional spreadsheets or being able to quickly learn new programs) as well as knowledge of basic computing concepts. We assigned respondents who selected only the statement "I don't have that much experience, but I can check my email and do simple tasks with word processing software" as an answer to the *novice* group.

▸ *Device and software usage.* When asked about the usage of different devices (see Fig. 5), 96.8% of the respondents affirmed that they use a smartphone in their daily lives and all respondents use either a laptop or a desktop computer on a daily basis.

When asked whether they use software systems more for work or personal reasons, 35.5% of the respondents stated that they use them more or less equally for work and personal reasons. 43.9% use software more for work (19.4% more often for work, 24.5% quite a bit more often for work) as shown in Fig. 6.

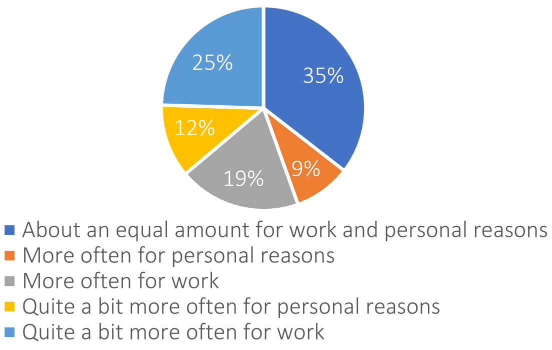


Fig. 6. Usage of software.

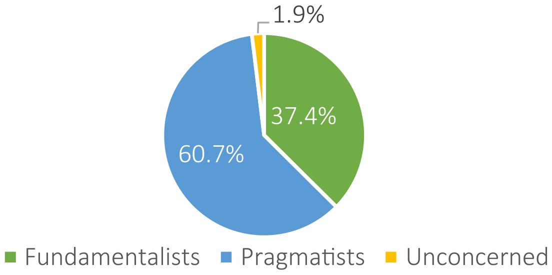


Fig. 7. PSI of respondents.

Both factors (usage of different devices and usage of software) evidence how software systems play an integral part in the everyday lives of the respondents.

4.2. RQ1 - privacy concerns

To determine respondents' concerns and worries respect to their privacy, we categorized respondents according to the PSI items, as described in Section 3.3.1. Subsequently, we quantified the level of risk our respondents perceive when sharing their information online (see risk beliefs, 3.3.2). Finally, we wanted to know what concerns the participants have regarding their privacy and what risks they see themselves exposed to.

4.2.1. Privacy segmentation index

Fig. 7 shows the distribution of the respondents according to the PSI items.

The distribution is largely consistent with data from Westin's privacy surveys (1996, 2000, 2001, and 2003) (Kumaraguru and Cranor, 2005), where the majority were classified as Privacy Pragmatists. The proportions of Privacy Fundamentalists and Privacy Unconcerned differ somewhat in our survey results. They show a higher proportion of Privacy Fundamentalists and a very low proportion of Privacy Unconcerned. Whether the Privacy Unconcerned have now become Pragmatists or Fundamentalists cannot be said on the basis of our data. The higher proportion of Privacy Fundamentalists could possibly be explained by the fact that people potentially value their (online) privacy more than it was the case in the past. The geographical distribution can also play a major role, since Germans are knowingly concerned about their online privacy (Schomakers et al., 2019). But here, too, it is not possible to prove this assumption only on the basis of our data.

▸ *Required Permissions and Installing Software.* Following the questions on the PSI, we asked respondents whether they pay attention to "required permissions" while installing apps. 60% indicate that they always pay attention, 34.2% do sometimes, and 5.8% do not pay attention. When asked how fast they press the "agree" button to terms and conditions when first using software, 30.3% of the respondents state that they press *instantly* the button, 55.5% *within one minute*, and 14.2% *spend more that one minute* before pressing the button.

50.0% of the Fundamentalists and 65.9% of the Pragmatists *always* pay attention to "required permissions". However, 66.6% (2 respondents) of the Privacy Unconcerned stated that they also always pay attention to them (the high percentage results from the fact that of the 155 respondents, a total of 3 were classified as Privacy Unconcerned). The proportion that *sometimes* pays attention is higher among Fundamentalists (41.3%) than among Pragmatists (29.8%).

When asked how fast they press "Agree" when installing software, the percentage who do so *within one minute*$^\diamond$ or *spend more than one minute** is also somewhat higher among Pragmatists (57.5%$^\diamond$, 14.9%*) than among Fundamentalists (53.5%$^\diamond$, 13.8%*), Privacy Unconcerned (33.3%$^\diamond$, 0%*).

Despite the fact that there is no significant difference between the groups, the results suggest that users are aware of privacy risks with respect to required permissions of apps because the majority pays attention to what permissions an app requires. When installing software, the majority also does not "agree" instantly. Our survey does not ask whether the users read information about data privacy during this time, for example. However, when this data is analyzed in conjunction with the question concerning needed permissions, we might infer that the respondents are well aware of the privacy threats that software may pose.

▸ *Privacy Policies.* Privacy policies are the primary channel through which service providers inform end-users about their data practices. 56.1% of the respondents *rarely* to *never* pay attention to whether a website provides a privacy policy, 23.9% *often* to *always* pay attention, and 20% pay attention *sometimes*. In fact, only 8.4% of respondents actually read a privacy policy (*often* to *always*), 13.5% said they *sometimes* read privacy policies and 78.1% *rarely* to *never*.

There is a large body of research that shows – as does our data – that privacy policies are not an appropriate medium for informing end-users about their privacy (Brunotte et al., 2022b; McDonald and Cranor, 2008; Reidenberg et al., 2015). Rather, they are made "by lawyers for lawyers". The fact that privacy policies are not read is not merely due to the users and their possible ignorance of their privacy (Karegar et al., 2020). Our results indicate that end users might be interested (Cummings et al., 2021) and concerned (Anton et al., 2010) about their online privacy. Thus, alternative techniques such as privacy explanations can be an adequate solution to foster transparency on data practices.

4.2.2. Risk beliefs

The calculated risk scores for the risk beliefs metric ranged from 1.95 to 7.0 (M = 4.78, SD = 1.19). The histogram depicted in Fig. 8 shows an approximately normal distribution for the risk scores. The Shapiro–Wilk test confirmed that the risk scores are normally distributed (W = 0.98, p = 0.058). The majority of the respondents (68.4%) have a risk score > 4.3. This suggests that respondents are not only aware of sharing their data online but also perceive it as a rather high risk.

We could observe that respondents with a higher risk score were more likely to read privacy policies with a positive correlation to Spearman's rank ($\rho = 0.41, p < 0.001$) and feel more uncomfortable using shopping portals ($\rho = 0.52, p < 0.001$), search tools ($\rho = 0.51, p < 0.001$), and games ($\rho = 0.45, p < 0.001$) (c.f. Fig. 6).

We also analyzed the possible relation between the IT experience level and the PSI. According to the PSI, the risk score among Privacy Fundamentalists is slightly higher on average (M = 5.52,

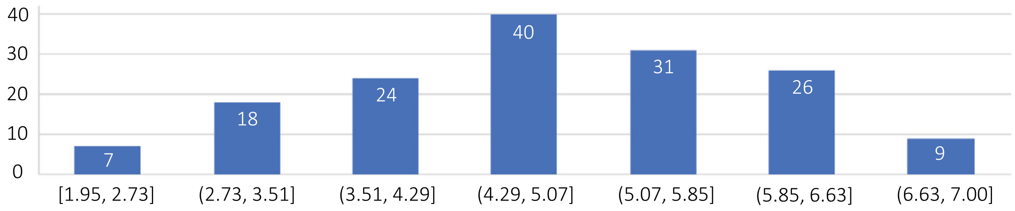


Fig. 8. Histogram of risk scores.

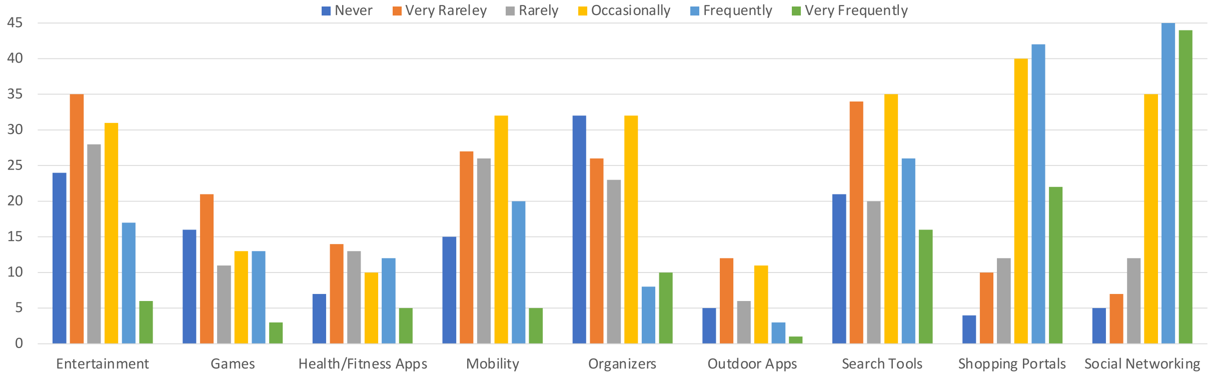


Fig. 9. How often do you feel uncomfortable about privacy when using software or visiting websites related to these categories?.

SD = 0.91) than for Pragmatists (M = 4.36, SD = 1.07). The risk score for Privacy Unconcerned (three respondents) is the lowest on average (M = 3.63, SD = 1.85). This difference between groups is statistically significant with a negative correlation according to Spearman's rank ($\rho = -0.49$, $p < 0.001$).

However, according to our findings, Intermediates (M = 5.01, SD = 1.11) and Novices (M = 5.13, SD = 0.94) have a higher average risk score than Experts (M = 4.55, SD = 1.14) and Developers (M = 4.77, SD = 1.34). Arguably, these differences are lower than between the PSI groups. Possibly, the lower risk score among the Experts and Developers could be due to the fact that they have a deeper understanding of software systems (especially the developers). Thus, they "have an idea of what software does internally". Intermediates and novices, on the other hand, see software systems primarily as black boxes (i.e., they know nothing about internal processes). This lack of knowledge could be the reason for the higher perceived risk. However, this assumption cannot be justified with our data since we did not ask any further questions in this direction.

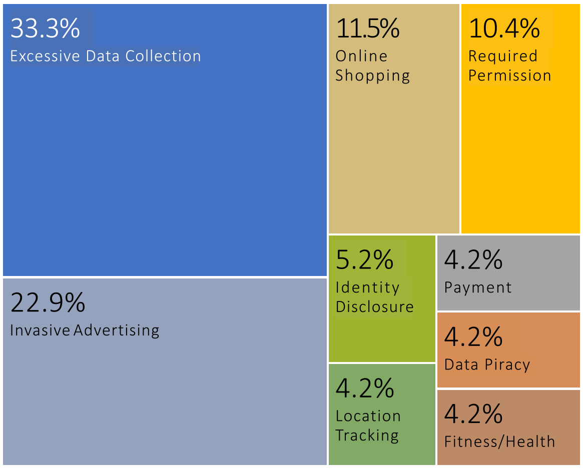


Fig. 10. Categories where respondents felt uncomfortable while using software.

4.2.3. *Privacy threats and concerns*

▶ *Using Software.* In order to assess what kind of privacy concerns and threats are faced by the respondents, we asked how often they feel uncomfortable about their privacy, depending on the use of different software. For this purpose, respondents previously selected which of the software categories they use at all. In this way, we ensured that the respondents could only indicate how uncomfortable they feel with software that they actually use. The results are depicted in Fig. 9.

▶ *Discomfort.* Respondents were asked to name a situation they had experienced in which they felt particularly uncomfortable using software in terms of privacy. By analyzing the respondents' answers, we were able to identify nine categories based on 96 codes. The categories are shown in Fig. 10.

The greatest discomfort among respondents is triggered by **excessive data collection** ("*it collects a lot of data about my everyday life*"), followed by **invasive advertising**. Invasive advertising is the practice of gathering and analyzing data and then presenting relevant ads based on that data. Respondents stated, for example, that they got the impression that offline conversations were being captured by technology such as smart-home assistants and that related advertising for products were being displayed during the next online search. Others expressed discomfort about advertising based on the chat history of messaging services.

Many respondents related that they feel uncomfortable shopping online because the shopping portals often collect a lot of data, such as address data, payment data, order history, and consumption behavior. In addition, several respondents complained that many permissions are often required to execute software (camera access, location, etc.) without being clear what these accesses are needed for.

▶ *Privacy Threats.* We also asked respondents for one threat they are particularly concerned about regarding their privacy (see Fig. 11). From the respondents' answers regarding their concerns, we were able to extract 171 codes. We then grouped these into seven categories. **Loss of control over data** is the concern most mentioned by respondents. Here, respondents expressed concern about not knowing who has access to their data, with whom this

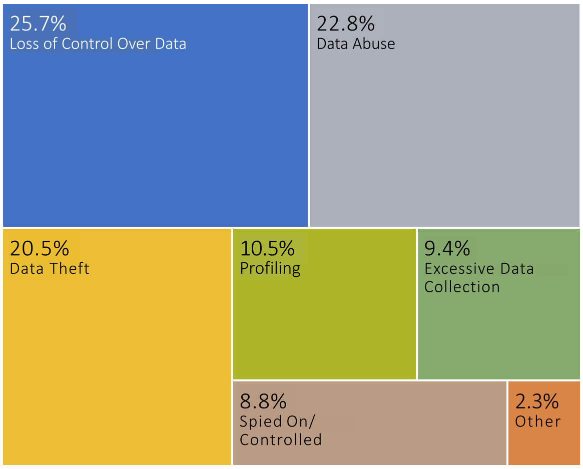


Fig. 11. Privacy threats that respondents are concerned about.

data is shared, and for what purpose this data is collected. Related concerns include **data theft** (*"I am afraid that they will steal my bank details"*, *"That one day our identities (like in the movies) can be stolen"*), **data abuse** (*"Use of the information to commit crimes"*, **profiling** (*"Creation and analysis of unique profiles via metadata consolidation"*). In addition, **excessive data collection** is also one of the concerns mentioned here, and some of the respondents are worried about being **spied on** (*"Software spies on me"*). In the category **Other**, we categorized statements that consisted of only one word (e.g., *"e-mail"*) or statements that would not fit into any of the other categories.

> **Answering RQ1:** Taking into account the risk scores (Section 4.2.2) and the concerns that our respondents expressed with respect to their privacy, we conclude that the majority of respondents have a high level of concern. Nevertheless, many respondents may be weighing the benefits and advantages of using certain services when it comes to their privacy. This could be supported by the PSI (60.65% Privacy Pragmatists) as well as the fact that respondents often have concerns about using social networking tools and shopping portals but at the same time, use them frequently on a daily basis.

4.3. RQ2 - current perception regarding privacy explanations

▸ *Interest in Explanations.* A hypothetical situation was presented to the respondents (Section 3). The goal was to analyze end users' need for privacy explanations in situations where a software system asks the user to disclose personal information. Based on this scenario, respondents were asked if they would be interested in a privacy explanation.

87.7% of the respondents are interested in receiving a privacy explanation (12.9% slightly interested, 20.6% moderately interested, 32.9% very interested, and 21.3% extremely interested), as shown in Fig. 12.

For the respondents who wanted to receive an explanation regarding their privacy (95.5%, $n = 148$), we presented an explanation of how their data would be used with respect to the given scenario. According to Spearman's rank, we found that respondents with a higher risk score were statistically significant more likely interested in receiving a privacy explanation ($\rho = 0.38, p < 0.001$). The privacy explanation regarding the use of the location was: *"in order to show you tours and recommendations near you, we need access to your location"* (E1). The explanation regarding the date of birth was: *"based on your date of birth, we can show you recommendations of what other users your age have liked"* (E2). Respondents were then asked how useful they found each of these privacy explanations and whether the explanations helped them feel more comfortable about disclosing personal information.

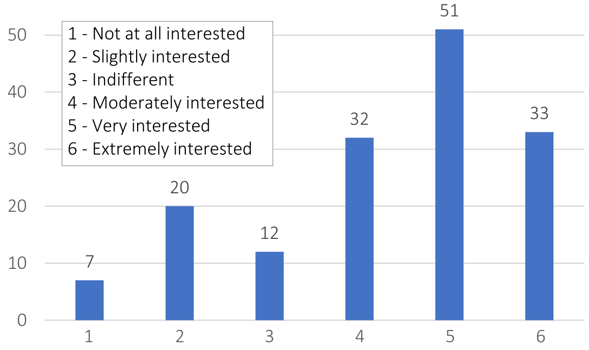


Fig. 12. How interested are you in receiving a privacy explanation? (hypothetical scenario).

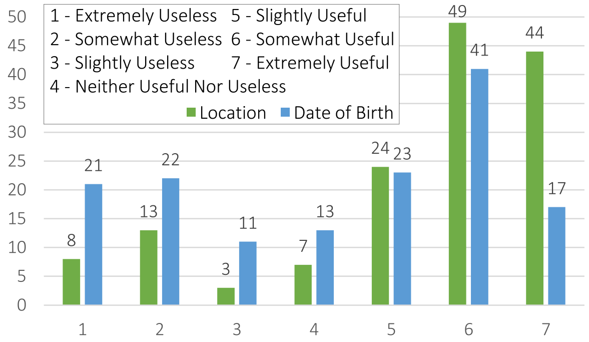


Fig. 13. How useful do you find these types of privacy explanations? ($n = 148$).

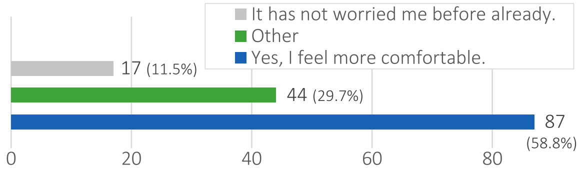


Fig. 14. Do you feel more comfortable with the privacy explanations? ($n = 148$).

▸ *Usefulness.* Fig. 13 shows how useful respondents ($n = 148$) found the privacy explanations. 79.1% indicated that they found the privacy explanation regarding the location (E1) useful (*slightly useful* to *extremely useful*) and 16.2% found them useless (slightly useless to extremely useless). The remaining 4.7% were indifferent. Regarding the date of birth (E2) (Fig. 13), 54.7% perceived the explanation useful and 36.5% found it was useless.

▸ *Well-being.* We asked if respondents felt more comfortable after receiving a privacy explanation. The results are shown in Fig. 14. Respondents who chose "other" as their answer had to justify their decision by entering an answer in the text box. The analysis of their answers resulted in 49 codes. We categorized the statements according to their meaning. The categories are **improvements** (14, 28.6%), **criticism** (4, 8.2%), **not sufficient** (14, 28.6%), **mistrust** (9, 18.4%), and **other** (8, 16.3%).

The group **improvements** included statements such as *"I would rather turn off the feature"* regarding the date of birth and *"classification in age group would be sufficient"*. In the category of **criticism**, statements such as *"interests do not depend on age"* were included. We did assign statements such as *"the reasoning*

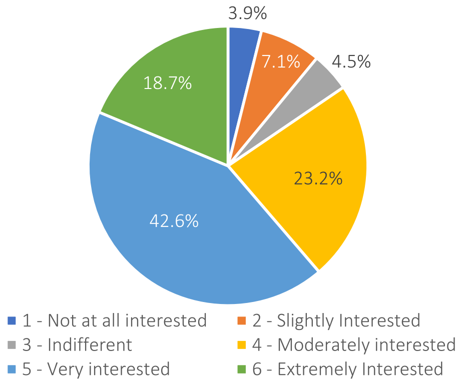


Fig. 15. Are you generally interest in privacy explanations?.

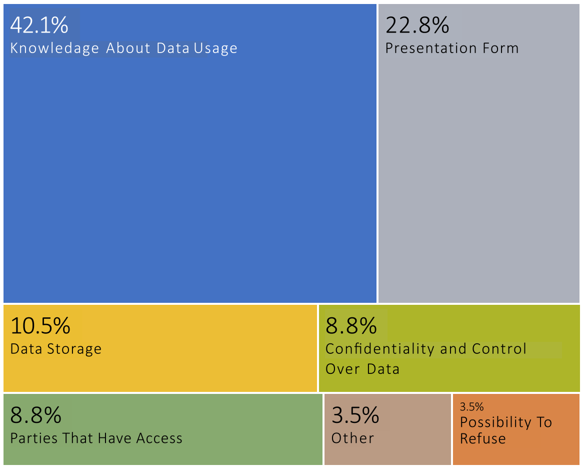


Fig. 16. Important aspects to be considered in privacy explanations.

on the age issue is not sufficient" as well as "*I would need to know in addition that these are the only reasons*" to the category **not sufficient**. Respondents' answers such as "*not trustworthy*" as well as "*feel spied on*" were assigned to the criticism category. Statements such as "*the declaration is nonsense. I didn't want it.*" have been assigned to the **other** category. In the scenario presented to the respondents, the app asked for the date of birth. We deliberately constructed the scenario in such a way that instead of the year of birth – which would have been technically sufficient for an age recommendation – the date of birth was requested. Seven of the respondents explicitly mentioned this.

▸ *General Interest.* When asked whether respondents are generally interested in receiving explanations with respect to their privacy, the majority (91.6%) of the respondents indicate that they are interested in privacy explanations (*extremely interested* to *slightly interested*), as shown in Fig. 15. When looking at respondents' risk score and their interest in receiving a privacy explanation, we get a similar picture as to the question from Fig. 12: a positive correlation according to Spearman's rank ($\rho = 0.29, p < 0.001$). The higher the risk Score, the statistically significant higher the interest in a privacy explanation.

▸ *Requirements on Privacy Explanations* In response to our open-ended question of what a privacy explanation should contain, respectively, what is expected of it, the analysis of the data resulted in 57 codes and revealed the following picture, as shown in Fig. 16.

First and foremost, end users want to understand how their data is being used. This includes knowledge about what data is used (42.1%), why, and how. The **presentation form** also plays an important role (22.8%). Respondents also indicated that explanations should be concise, precise, and written in simple language. Additionally, respondents said that icons can be visually supportive.

Respondents also expressed that they would like to know where the data is stored, for how long, how it is protected, and if or when the data is deleted. We clustered statements of this type in the **data storage** (10.5%) group. Respondents want to "*be sure that this data cannot be sold to other companies*". They want to be assured that the data will be kept confidential and "*not used for anything else*" (**confidentiality and control over data**, 8.8%). Respondents also indicated that privacy explanations should provide information about **parties that have access** (8.8%) to the data and give information about a **possibility to refuse** (3.5%). The category **other** comprises statements for which we were not able to make a relation to our question.

> **Answering RQ2:** The vast majority of respondents (91.6%) are interested in privacy explanations and consider them useful since they inform them about data practices. For this purpose, it is important for them to know what the data is used for and how it is stored. The presentation form of such an explanation also plays an important role. A privacy explanation should be easy to understand and be connected to the user's present context. That means, for example, if an app requires a user's location, the app should explain why the location is needed.

4.4. RQ3 - privacy explanations and the concept of trust

To answer RQ3, we asked respondents to name up to three benefits they think may be associated with privacy explanations. Following this, we asked when they should be presented and whether respondents agree that privacy explanations can help increase the level of trust in a software system.

▸ *Benefits.* 137 respondents answered this open-ended question, which resulted in 135 valid responses. Each valid response was analyzed and resulted in a total of 363 codes.

We summarized 24 codes (6.6%) in the category **other**, which we could not assign to any advantage. These statements could not have been assigned to a disadvantage either, because they partly showed a lack of understanding on the part of the respondents ("*data sale*", "*minimizes spam*", "*the company makes money with it*"). Since we would have to interpret too much into such statements and a classification into another category would be too subjective, we decided to classify such statements into this cluster.

Transparency is the largest category and 39.4% (143) of the 363 codes were grouped here. Statements such as "*transparency with the user*", "*clarity*" or "*increases users understanding of the software they are using*" were included in this category.

The second largest cluster is **foster trust** (47 codes, 12.9%). Respondents mentioned benefits such as "*strengthen trust*", "*build trust towards the system*", and "*more trust in the application and the company*". In addition to trust-fostering benefits of privacy explanations, respondents expressed that explanations can also create a **sense of well-being** (40, 11.0%). In this category, we have grouped statements such as "*reduces uncertainty*", "*transmits a sense of confidence*", and "*improves users' sense of well-being*".

In addition, according to the respondents, privacy explanations contribute to make **conscious choices** (17, 4.7%) when disclosing personal data. They state that an explanation "*allows users to give informed consent*", or "*one can decide more consciously whether it is worthwhile to disclose the data*". According to the respondents, **self-determination** (10, 2.8%) might be a further benefit of privacy explanations because it gives "*control over your own data*".

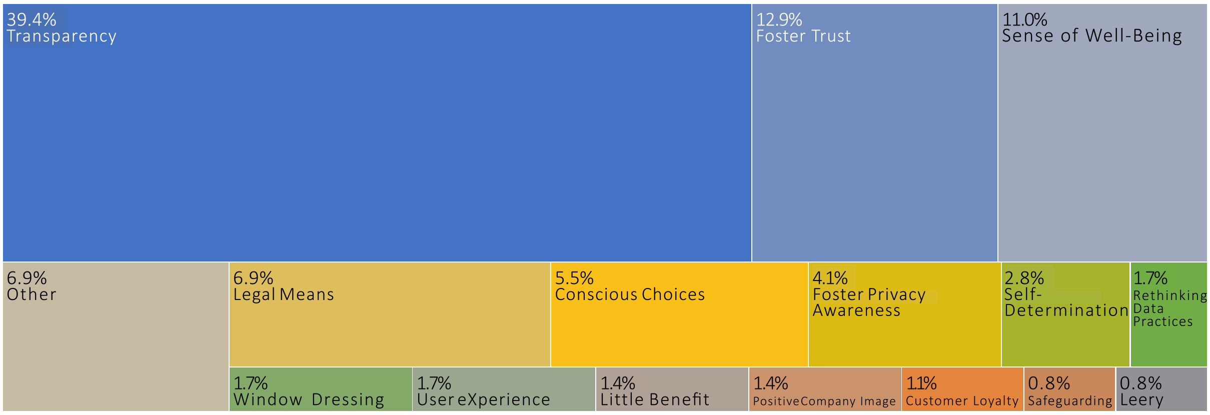


Fig. 17. What do you think is the benefit of privacy explanations?.

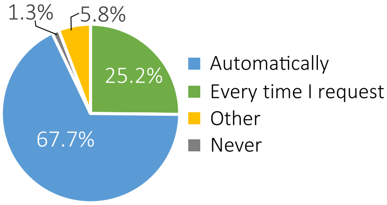


Fig. 18. When should a privacy explanation be presented?.

We have grouped statements such as *"helps people to think twice"* as well as *"becoming more aware of what happens to data"* in the category **foster privacy awareness** (15, 4.1%).

In addition to these categories, respondents also see a benefit from a legal perspective. Statements such as *"comply with Federal legislation"* or *"compliance"* were grouped into the category **legal means** (25, 6.9%). Other benefits of privacy explanations considered by respondents include a **positive company image** (5, 1.4%), **customer loyalty** (4, 1.1%), and a **rethinking of current data practices** (6, 1.7%). Here they expressed privacy explanations *"force a company to think about what data to collect"* or *"the operators inflict a usage policy upon themselves, which sets a boundary"*.

In addition, respondents indicate that privacy explanations foster **user experience** (UX) (6, 1.7%) as well as that they may safeguard a person's interests (*"protection of own interests"*) or the person themselves (*"protection of the own person from, e.g., unnecessary advertising"*). We coded statements like the last two as **safeguarding** (3, 0.8%).

While some of the respondents see only **little benefits** (5, 1.4%), others also expressed criticism. They described privacy explanations as **window dressing** since they could (*"fool the users"* and may cause *"Consumer confusion (appeasement/downplay)"*). Other statements suggest that respondents are rather leery of privacy explanations. We have therefore assigned statements such as *"you can not trust that this will be adhered to"* to the **leery** (3, 0.8%) category.

▸ *When to Present.* We asked respondents when privacy explanations should be presented. The results are shown in Fig. 18.

The analysis of these answers resulted in 18 codes. We have divided the statements into the following categories according to the answer options *automatically* (6, 31.5%), *on change* (6, 31.5%), and *on request* (6, 31.5%). 5.8% of respondents chose *other* as an answer option and provided their answer in text form. The analysis of the answers reveal that respondents want an explanation presented automatically, and every time something changes, but also whenever they request an explanation.

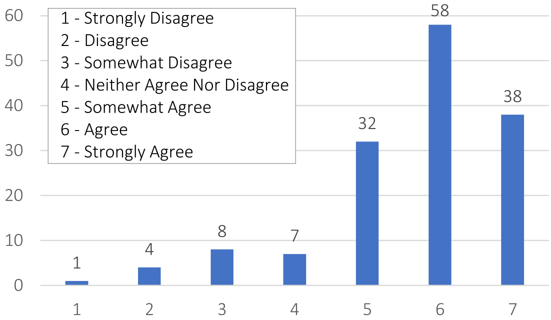


Fig. 19. Can privacy explanations be a possible factor to increase the level of trust in a software system? ($n = 148$).

▸ *Privacy Explanations and Trust.* We asked our respondents ($n = 148$) whether they agree that privacy explanations can be a possible factor to increase the level of trust in a software system. We received a clear picture here, as shown in Fig. 19, which is in line with responses from Fig. 17.

> **Answering RQ3:** Our results suggest that privacy explanations might foster trust as well as privacy awareness. They can provide more transparency with respect to the applied data practices of a system and enable end users to make more conscious choices regarding their privacy (Fig. 17).
> Privacy explanations can make end users feel more secure and increase their well-being while operating a software system. Furthermore, privacy explanations might have a positive impact on the trustworthiness of a software system and, in turn, foster end user trust in the system.

5. Discussion

Our results led to six findings that helped to shape four concepts. In this section, we discuss these concepts and the related findings. The first concept explores the end users' perception on privacy explanations and the privacy paradox, discussed in Section 5.1. The second concept explores the relationship between trust and privacy explanations and introduces the trust-coin metaphor, discussed in Section 5.2. The third concept explores the requirements on privacy explanations, discussed in

Section 5.3. And the fourth concept explores the relationship between relevant stakeholders for privacy-aware systems, discussed in Section 5.4.

5.1. Users' perception on privacy explanations

Our findings show that participants' perceptions of privacy explanations are paradoxical: many participants do not actively take action about data practices, despite the fact that they see privacy and privacy explanations as desirable and beneficial. Two findings enable us to more fully grasp the subtleties of the potential end users' perception of privacy explanations:

▸ *Finding 1: Privacy explanations are beneficial.* According to our results, a vast majority (91.6%) of the respondents are generally interested in privacy explanations. A closer look at the explanations given in the hypothetical scenario and the respondents' reactions to those explanations reveal that they perceive privacy explanations as supportive and feel more comfortable (Fig. 14) when they receive information about data usage.

74.3% (58.8% with respect to E2) of the respondents who received an explanation felt more comfortable after knowing about how personal data would be used. Respondents indicate that explanations reduce their uncertainty, provide more security and confidence, which all, in turn, result in a feeling of well-being.

Informing the user *(a)* that personal data about them is being collected, *(b)* what data is being collected, and *(c)* how this data is being used, contributes to privacy awareness, as described in Pötzsch (2009). At the same time, being informed enables self-determined and conscious decisions when using software systems, which was also mentioned by the respondents. Following this line of thought, it also becomes clear why some of the respondents consider privacy explanations as a kind of *safeguard*. A safeguard between one's own *online privacy* and the uninformed/unintentional disclosure of privacy aspects.

▸ *Finding 2: The privacy paradox.* To evaluate the privacy attitudes of our respondents, we grouped them according to the PSI items (Section 4.2.1), calculated their risk beliefs (Section 4.2.2), and asked them what concerns they have in terms of privacy (Section 4.2.3).

Even though respondents are concerned and care about their privacy they avoid engaging in the necessary privacy behavior. This phenomenon is known in the literature as **privacy paradox**. This term was coined by Barnes (2006) and is well researched by many others (Gerber et al., 2018; Kokolakis, 2017; Hargittai and Marwick, 2016; Bandara et al., 2020; Pötzsch, 2009). In a nutshell, the privacy paradox states, "I am aware that my privacy is being violated, yet I continue to utilize this service". The privacy paradox affects all generations (Pentina et al., 2016) and "cannot be attributed solely to either a lack of understanding of or a lack of interest in privacy" (Hargittai and Marwick, 2016), as reflected in our results.

Furthermore, according to Hargittai and Marwick (2016), end users' lack of appropriate privacy behavior is rooted in the apathy they developed towards online privacy since the systems are often black boxes, have opaque data practices, and the privacy controls available change frequently. As a result, it is difficult and confusing for end users to comprehend how their personal data flows. This leads to frustration and worries, which yields self-censorship and apathy. Therefore, easy accessible information should be available to inform end users about data practices.

This paradoxical attitude of end users (Gerber et al., 2018; Mourey and Waldman, 2020) is one aspect that conveys the complexity and challenges of dealing with privacy and is also reflected in our results. This complexity stems from the fact that users weigh costs and benefits when making decisions (Earp and Baumer, 2003). Different user attitudes also reflect on privacy behavior (Bates, 1964; Rudolph et al., 2018). For instance, some users are more concerned or "leerier" than others. This phenomenon was also evident when we asked about the benefits of privacy explanations, where some of the respondents expressed their concerns and worries. These concerns also reflect an important point related to privacy: *trustworthiness*. The next Section 5.2 will explore trust and trustworthiness in more detail.

5.2. The trust-coin—trust and trustworthiness

Trust in IT is an important concept since today's society heavily relies on IT. Trust in the competence of a system means that a system has the functionality or functional capability to perform a particular task that the trustor requests (McKnight, 2005). In terms of privacy, opaque, incomprehensible data practices, and lack of transparency can harm this trust. Our participants' responses also confirmed this.

To better understand trust and trustworthiness, let us imagine trust as a coin: a "trust-coin". On the one side of the coin is the end-user trust, which represents their perception of trust towards a system. On the other side of the coin is trustworthiness as a property or quality aspect of a system. In terms of privacy, we cannot look at the two sides separately because they are interwoven. This means that if software engineers want end users to trust their system, they should ensure that their system *is* trustworthy in terms of privacy.

▸ *Finding 3: Privacy explanations as a means to trust.* According to our results, privacy explanations are a means to increase the level of trust in software systems (Fig. 19). Our results suggest that they might help to establish a relationship of trust between the end user and the system by increasing data transparency and clarity. Privacy explanations might help to put the user in control so that they can make self-determined, conscious choices with respect to their personal data. Respondents' answers reveal what requirements on privacy explanations (Fig. 16) can serve respondents' privacy concerns.

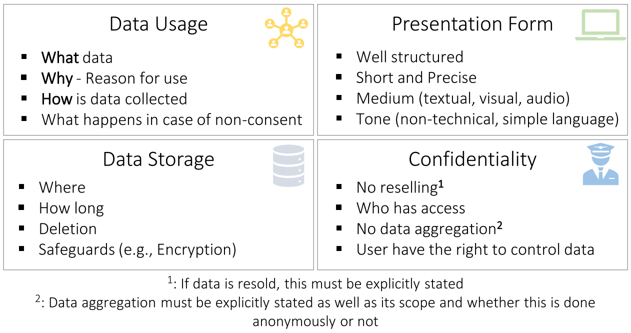
5.3. Requirements for privacy explanations

To incorporate privacy explanations into systems, stakeholder requirements need to be elicited. It is important to meet the needs and expectations that stakeholders have regarding such explanations, as evidenced in *Finding 1*. Otherwise, they may defeat their purpose of informing the end user regarding their privacy or even cause mistrust (Papenmeier et al., 2019; Pieters, 2011).

In light of this, it is not enough to provide "any" explanation. A privacy explanation must make sense to the end user, considering their individual characteristics and context (c.f. explanation E1 and E2 in Fig. 13).

This provides important insights for the development of explainable and privacy-aware systems. End users demand data reduction and data economy because their privacy is important to them, as mentioned above. That means for the design of the systems, data economy, responsible, and fair use of personal data is required (Koskinen et al., 2019; Schafer and Edwards, 2017). For our scenario, this means that instead of asking for the date of birth, this system should only ask for the year of birth. Taking into account the respondents' suggestions for improvement, an explanation should be able to provide further information on demand. For example, what the consequences are for the user if they provide their consent.

▸ *Finding 4: The "trust" side of the coin.* We asked respondents what they expect from a privacy explanation or what it should contain (Fig. 16). This helped us to identify aspects that should be considered in privacy explanations, according to the



Data Usage

- **What** data
- **Why** - Reason for use
- **How** is data collected
- What happens in case of non-consent

Presentation Form

- Well structured
- Short and Precise
- Medium (textual, visual, audio)
- Tone (non-technical, simple language)

Data Storage

- Where
- How long
- Deletion
- Safeguards (e.g., Encryption)

Confidentiality

- No reselling[1]
- Who has access
- No data aggregation[2]
- User have the right to control data

[1]: If data is resold, this must be explicitly stated

[2]: Data aggregation must be explicitly stated as well as its scope and whether this is done anonymously or not

Fig. 20. Requirements for privacy explanations.

answers in our survey and in line with our definition of online privacy (Section 2.2). These aspects can work as high-level requirements for privacy explanations that should be met, as depicted in Fig. 20. We categorized these high-level requirements into four categories. These first requirements in conjunction with our definition of online privacy can serve as a starting point for software engineers to understand what elements should be considered when designing privacy explanations. We list each one of the four categories below.

Data Usage: Information about the use of personal data is crucial to clarify **what** data is used, **why**, and **how** it is used. We refer to this as the **2W1H** principle. A privacy explanation should inform the user about it.

Data Storage: Users should know where and how long the data is stored. An explanation should provide information whether the data is temporarily collected or in the long-run, as well as what **safeguards** are put in place to prevent accidental or deliberate privacy violations.

Confidentiality: Confidentiality is also an essential point in terms of privacy. This includes disclosing who has access to the data and whether the data is resold or used for further aggregation. This includes opt-out mechanisms (e.g., for reselling data) to ensure that users retain control over their own data.

Presentation Form: Privacy explanations should be well-structured, short, and precise, as well as communicated in a non-technical and simple language.

These requirements also reflect what we claim in our proposed definition of online privacy (Section 2.2). What privacy aspects an individual is willing to share with others corresponds to the data usage and confidentially categories. The technical implementation of these categories is realized via the requirements for presentation form and data storage.

Privacy explanations should be seen as a **context-aware, usable privacy feature**. According to our findings, a system should provide such explanations in the relevant context (automatically and/or on request when certain personal data are used) and be able to react and (re-)enter into a dialog with the end user, in case a policy changes.

▸ *Finding 5: The "trustworthy" side of the coin.* In order to gain the trust of end users, systems must be reliable, i.e., trustworthy in terms of privacy. Therefore, systems that encompass quality aspects such as accountability, fairness, and ethics (Koskinen et al., 2019; Rantanen, 2019) are needed as well as where privacy is the "default setting" (Cavoukian et al., 2009). In addition, these systems must implement applicable laws and regulations in terms of privacy in order "to protect individual privacy in information processing" (Bowman et al., 2015).

With regard to privacy explanations, this means if a system explains to a user for what purpose a certain privacy aspect is needed, the system must not use it for any other purposes. The system must guarantee this in order to count as trustworthy. By giving explanations, the system can support the end user in building trust in the system itself. This also illustrates why it is important not only to distinguish between trust and trustworthiness, but also to specifically take both quality aspects into account.

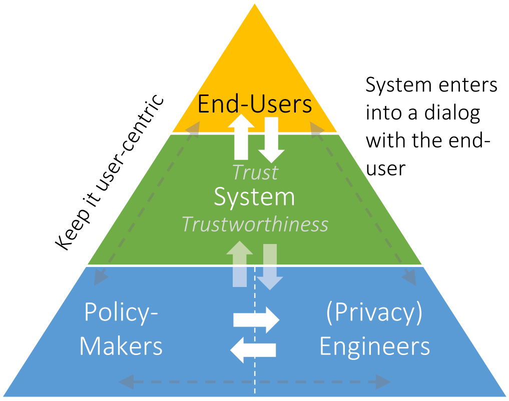


Fig. 21. Trialog of privacy.

5.4. Trialog of privacy – Keep it user-centric

"Privacy requires a dialogue between two types of people: those who speak policy and those who speak engineering" (Ohm, 2015). This means that policy-makers (e.g., legislators and lawyers) must work together in direct dialog with software engineers and support each other in eliciting privacy requirements, assuring that they are legally compliant, and translating them into systems, regulations, and norms. According to Ohm, lawyers respond and react to what engineers engineer instead of communicating with each other (Ohm, 2015).

From our point of view, the user, who should actually be the focus (be the center), takes a back seat in Ohm's statement. We propose a **trialog of privacy** (see Fig. 21). Therefore, we enhance Ohm's statement and suggest:

> Privacy requires a trialog between three types of people: those who are end users and whose privacy is at stake, those who speak (privacy) engineering, and those who speak policy.

▸ *Finding 6: Relevant stakeholders for privacy-aware systems.* According to Fig. 21, the foundation for a privacy-aware system lies in the dialog between policy-makers and privacy engineers. We adopt the term privacy engineer in reference to Bowman et al. (2015) who stated that *privacy is not something that can be fully addressed with a few architectural decisions made in the design phase alone*. The commitment to privacy is dynamic. As technology grows and is adopted by more users in different contexts, this commitment needs to be maintained. Therefore, Bowman postulates the role of a privacy engineer who maintains and is responsible for the privacy architecture (Bowman et al., 2015). Similar to the role of a usability engineer who is responsible for the usability architecture.

When these two parties (policy-makers and privacy engineers) are in dialog, the system might be built on a privacy-aware and trustworthy basis. To consider the end user as a third party and complement the trialog the system should "consciously designed around the interests and needs of individual users" (Cavoukian et al., 2009) in order to meet their individual privacy preferences and enter into a dialog with the users via explanations.

This is what is meant by **keep it user-centric**[2] and why these three parties should be counted among the group of relevant stakeholders.

6. Threats to validity

The strategy to select the participants has some limitations. Despite the fact that we received answers from different countries of the world, the majority of responses came from Germany and Brazil. This may not reflect the whole population and may threaten the global generalizability of our results. Although 155 participants provided a substantial amount of responses, some of the conclusions might be affected by this size and should not be overgeneralized. Most of the respondents of our study have profound IT knowledge. Our population may not take into account people who have difficulties operating software systems. Therefore, we cannot generalize the needs regarding privacy explanations, but we get an overview of what different people think. For RQ2 we identify different concerns while using applications. We only can evaluate the answers of our respondents and it might be that there are much more reasons of concerns for privacy, e.g. people who do not have IT knowledge. The other findings for our research have the same limitations. To find more concerns further experiments need to be conducted. To mitigate the thread that the analysis is too subjective we use *in vivo coding* by two researchers. Each of them categorizes the data. During the second cycle they discuss and compare their findings to increase the consistency and reliability.

To evaluate the need for privacy explanation we use a hypothetical scenario. This scenario is potentially not the daily use for the participants, but it might encounter them in real life. Only users of smartphones could better empathize with this situation. This situation confronted the respondents only in a scenario of vacation. The results could be different when they will confront in a business or financial scenario, because it could have posed a greater threat to their privacy.

Another aspect is that a good question wording and instrumentation layout are crucial for the results of a survey. We followed guidelines and conducted pilot-tests to ensure these aspects. However, the order of questions in the questionnaire may have impacted in the participants' understanding about whether we were asking questions about the need to receive explanations in a general context or related to the previous question about a more specific context. However, we considered that this would be helpful for participants who could have difficulties to imagine other situations where they would need explanations.

We decided to disclose our online questionnaire and raw data (Brunotte, 2022) so that other researchers may be able to replicate and comprehend how we have drawn our conclusions and recommendations from the data. This step should serve as a final strategy to mitigate threats to the internal validity.

7. Future directions

Our study provided us with valuable insights into how respondents perceive privacy explanations and that these can make an important contribution to communicating data practices to end users in an comprehensible and transparent way.

Building upon findings of our research, we need to investigate how to translate our set of requirements for privacy explanations into a system. To keep privacy explanations *as simple as possible* in order not to overwhelm the user, we plan a user study in which we survey how privacy explanation must be engineered (Brunotte et al., 2022a). We assume, that an layered information structure may be beneficial. In the first place, a system informs a user about personal information usage, as happened in our scenario. In a next step, the system shall provide further information to the end users upon request, according to their individual privacy preferences and our 2W1H principle. Overall, it is important that it follows an actionable and operationalizable process.

Finally, we suggest more collaboration across disciplines (humanities, law, and computer science) since "privacy is not an individual process, but rather a collective effort that requires cooperation" (Hargittai and Marwick, 2016) of those who are involved. Research could address how the trialog of privacy could be integrated into existing privacy-frameworks (Senarath et al., 2017; Notario et al., 2015) that are currently focused on the engineering part.

8. Conclusion

In this article, we conducted an online survey with 155 participants to investigate end users perception and attitude towards privacy explanations.

Our findings suggest, that end users perceive privacy explanations as beneficial and they may influence the well-being of end users. 91.6% of our respondents are generally interested in receiving such explanations. We found that privacy explanations may also be seen as a means toward end user trust and aim to bridge the gap in information asymmetry between end users and software systems by mitigating opacity with respect to data practices and thus supporting end users in making more conscious choices.

The results of this study expanded our understanding of how the individual privacy preferences of users might be retained. We were able to derive a set of high-level requirements for privacy explanations. These can serve as a starting point for software engineers to incorporate privacy explanations in software systems. We conclude that the integration of privacy explanations needs to be conducted carefully to meet the different users' requirements on privacy explanations.

Furthermore, we propose the trialog of privacy approach as a paradigm for the development of privacy-aware systems since more interdisciplinary collaboration is needed in order to address the complex challenges that arise in terms of privacy.

CRediT authorship contribution statement

Wasja Brunotte: Conceptualization, Methodology, Validation, Investigation, Formal analysis, Writing – original draft, Writing – review & editing, Visualization, Resources. **Alexander Specht:** Writing – original draft, Writing – review & editing, Conceptualization, Methodology, Validation, Investigation, Data curation, Resources, Formal analysis. **Larissa Chazette:** Conceptualization, Methodology, Validation, Investigation, Writing – original draft, Writing – review & editing. **Kurt Schneider:** Supervision, Conceptualization, Methodology, Validation, Writing – review & editing, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Our online questionnaire and raw data is available at Brunotte (2022).

[2] *Respect for User Privacy – Keep it User-Centric:* 7th Foundational Principle of Privacy by Design by Cavoukian et al. (2009).

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy within the Cluster of Excellence PhoenixD (EXC 2122, Project ID 390833453) and by the research initiative Mobilise between the Technical University of Braunschweig and Leibniz University Hannover, funded by the Ministry for Science and Culture of Lower Saxony, Germany.

References

Allen, A.L., 1988. Uneasy Access: Privacy for Women in a Free Society. Rowman & Littlefield.

Amparore, E., Perotti, A., Bajardi, P., 2021. To trust or not to trust an explanation: using LEAF to evaluate local linear XAI methods. PeerJ Comput. Sci. 7, e479.

Anton, A.I., Earp, J.B., Young, J.D., 2010. How internet users' privacy concerns have evolved since 2002. IEEE Secur. Priv. 8 (1), 21–27. http://dx.doi.org/10.1109/MSP.2010.38.

Balkir, E., Kiritchenko, S., Nejadgholi, I., Fraser, K.C., 2022. Challenges in applying explainability methods to improve the fairness of NLP models. arXiv:2206.03945.

Bandara, R., Fernando, M., Akter, S., 2020. Explicating the privacy paradox: A qualitative inquiry of online shopping consumers. J. Retail. Consum. Serv. 52, 101947. http://dx.doi.org/10.1016/j.jretconser.2019.101947.

Barnes, S.B., 2006. A privacy paradox: Social networking in the United States. First Monday 11 (9), http://dx.doi.org/10.5210/fm.v11i9.1394.

Barnett White, T., 2004. Consumer disclosure and disclosure avoidance: A motivational framework. J. Consum. Psychol. 14 (1), 41–51. http://dx.doi.org/10.1207/s15327663jcp1401&2_6.

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., Herrera, F., 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Inf. Fusion 58, 82–115. http://dx.doi.org/10.1016/j.inffus.2019.12.012.

Bates, A.P., 1964. Privacy — A useful concept? Soc. Forces 42 (4), 429–434. http://dx.doi.org/10.2307/2574986.

Bhave, D.P., Teo, L.H., Dalal, R.S., 2020. Privacy at work: A review and a research agenda for a contested terrain. J. Manag. 46 (1), 127–164. http://dx.doi.org/10.1177/0149206319878254.

Bloustein, E.J., 1964. Privacy as an aspect of human dignity: An answer to dean prosser. N. Y. Univ. Law Rev. 39 (6), 962–1007.

Bowman, C., Gesher, A., Grant, J.K., Slate, D., 2015. The Architecture of Privacy, first ed. O'Reilly Media, Inc., Sebastopol, CA, USA.

Bräunlich, K., Dienlin, T., Eichenhofer, J., Helm, P., Trepte, S., Grimm, R., Seubert, S., Gusy, C., 2021. Linking loose ends: An interdisciplinary privacy and communication model. New Media Soc. 23 (6), 1443–1464. http://dx.doi.org/10.1177/1461444820905045.

Brunotte, W., 2022. Data for research article "Privacy Explanations – A Means to End-User Trust". http://dx.doi.org/10.5281/zenodo.7215560.

Brunotte, W., Chazette, L., Klös, V., Speith, T., 2022a. Quo vadis, explainability? – A research roadmap for explainability engineering. In: Gervasi, V., Vogelsang, A. (Eds.), Requirements Engineering: Foundation for Software Quality. Springer International Publishing, Cham, pp. 26–32. http://dx.doi.org/10.1007/978-3-030-98464-9_3.

Brunotte, W., Chazette, L., Kohler, L., Klunder, J., Schneider, K., 2022b. What about my privacy? Helping users understand online privacy policies. In: Proceedings of the International Conference on Software and System Processes and International Conference on Global Software Engineering. ICSSP '22, Association for Computing Machinery, New York, NY, USA, pp. 56–65. http://dx.doi.org/10.1145/3529320.3529327.

Brunotte, W., Chazette, L., Korte, K., 2021. Can explanations support privacy awareness? A research roadmap. In: 2021 IEEE 29th International Requirements Engineering Conference Workshops (REW). pp. 176–180. http://dx.doi.org/10.1109/REW53955.2021.00032.

Cavoukian, A., et al., 2009. Privacy by design: The 7 foundational principles. In: Information and Privacy Commissioner of Ontario, Canada, Vol. 5. p. 12.

Chakraborti, T., Sreedharan, S., Grover, S., Kambhampati, S., 2019. Plan explanations as model reconciliation – An empirical study. In: 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI). pp. 258–266. http://dx.doi.org/10.1109/HRI.2019.8673193.

Chang, C., Li, H., Zhang, Y., Du, S., Cao, H., Zhu, H., 2019. Automated and personalized privacy policy extraction under GDPR consideration. In: Biagioni, E.S., Zheng, Y., Cheng, S. (Eds.), Wireless Algorithms, Systems, and Applications. Springer International Publishing, Cham, pp. 43–54.

Charmaz, K., 2014. Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis, second ed. SAGE Publications Inc., Thousand Oaks, CA, USA.

Chazette, L., Brunotte, W., Speith, T., 2021. Exploring explainability: A definition, a model, and a knowledge catalogue. In: 2021 IEEE 29th International Requirements Engineering Conference (RE). pp. 197–208. http://dx.doi.org/10.1109/RE51729.2021.00025.

Chazette, L., Karras, O., Schneider, K., 2019. Do end-users want explanations? Analyzing the role of explainability as an emerging aspect of non-functional requirements. In: 2019 IEEE 27th International Requirements Engineering Conference (RE). pp. 223–233. http://dx.doi.org/10.1109/RE.2019.00032.

Chazette, L., Schneider, K., 2020. Explainability as a non-functional requirement: challenges and recommendations. Requir. Eng. 25 (4), 493–514. http://dx.doi.org/10.1007/s00766-020-00333-1.

Cohen, J., 1968. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. Psychol. Bull. 70 (4).

Consolvo, S., Smith, I.E., Matthews, T., LaMarca, A., Tabert, J., Powledge, P., 2005. Location disclosure to social relations: Why, when, & what people want to share. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '05, Association for Computing Machinery, New York, NY, USA, pp. 81–90. http://dx.doi.org/10.1145/1054972.1054985.

Cronbach, L.J., 1951. Coefficient alpha and the internal structure of tests. Psychometrika 16 (3), 297–334. http://dx.doi.org/10.1007/BF02310555.

Cummings, R., Kaptchuk, G., Redmiles, E.M., 2021. "I need a better description": An investigation into user expectations for differential privacy. In: Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security. CCS '21, Association for Computing Machinery, New York, NY, USA, pp. 3037–3052. http://dx.doi.org/10.1145/3460120.3485252.

Dahl, E.S., 2018. Appraising black-boxed technology: the positive prospects. Philos. Technol. 31 (4), 571–591. http://dx.doi.org/10.1007/s13347-017-0275-1.

Dai, E., Zhao, T., Zhu, H., Xu, J., Guo, Z., Liu, H., Tang, J., Wang, S., 2022. A comprehensive survey on trustworthy graph neural networks: Privacy, robustness, fairness, and explainability. http://dx.doi.org/10.48550/ARXIV.2204.08570.

De Terwangne, C., 2012. Internet privacy and the right to be forgotten/right to oblivion. In: VII Congreso Internacional Internet, Derecho y Política. Neutralidad de la Red y Otros Retos Para El Futuro de Internet,[Monografía Online], IDP, Revista de Internet, Derecho y Política, UOC. In: VII, no. 13, pp. 109–121.

Dinev, T., 2014. Why would we care about privacy? Eur. J. Inf. Syst. 23 (2), 97–102. http://dx.doi.org/10.1057/ejis.2014.1.

Earp, J., Anton, A., Aiman-Smith, L., Stufflebeam, W., 2005. Examining internet privacy policies within the context of user privacy values. IEEE Trans. Eng. Manage. 52 (2), 227–237. http://dx.doi.org/10.1109/TEM.2005.844927.

Earp, J.B., Baumer, D., 2003. Innovative web use to learn about consumer behavior and online privacy. Commun. ACM 46 (4), 81–83. http://dx.doi.org/10.1145/641205.641209.

Ehsan, U., Tambwekar, P., Chan, L., Harrison, B., Riedl, M.O., 2019. Automated rationale generation: a technique for explainable AI and its effects on human perceptions. In: Proceedings of the 24th International Conference on Intelligent User Interfaces. pp. 263–274.

Elahi, H., Castiglione, A., Wang, G., Geman, O., 2021. A human-centered artificial intelligence approach for privacy protection of elderly App users in smart cities. Neurocomputing 444, 189–202. http://dx.doi.org/10.1016/j.neucom.2020.06.149, URL https://www.sciencedirect.com/science/article/pii/S0925231221001259.

Elahi, G., Yu, E., 2009. Trust trade-off analysis for security requirements engineering. In: 2009 17th IEEE International Requirements Engineering Conference. pp. 243–248. http://dx.doi.org/10.1109/RE.2009.12.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., Vayena, E., 2018. AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds Mach. 28 (4), 689–707. http://dx.doi.org/10.1007/s11023-018-9482-5.

Garcia-Rivadulla, S., 2016. Personalization vs. privacy: An inevitable trade-off? IFLA J. 42 (3), 227–238. http://dx.doi.org/10.1177/0340035216662890.

George, D., Mallery, P., 2009. SPSS for Windows Step by Step: A Simple Study Guide and Reference, 17.0 Update, tenth ed. Allyn & Bacon, Inc., USA.

Gerber, N., Gerber, P., Volkamer, M., 2018. Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. Comput. Secur. 77, 226–261. http://dx.doi.org/10.1016/j.cose.2018.04.002.

Giorgini, P., Massacci, F., Mylopoulos, J., Zannone, N., 2004. Requirements engineering meets trust management. In: Jensen, C., Poslad, S., Dimitrakos, T. (Eds.), Trust Management. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 176–190.

Groves, R.M., Fowler, Jr., F.J., Couper, M.P., Lepkowski, J.M., Singer, E., Tourangeau, R., 2011. Survey Methodology, second ed. John Wiley & Sons, Hoboken, NJ, USA.

Hann, I.-H., Hui, K.-L., Lee, T., Png, I., 2002. Online information privacy: Measuring the cost-benefit trade-off. In: International Conference on Information Systems (ICIS). pp. 1–11.

Harari, Y.N., 2015. Sapiens: A Brief History of Humankind, first ed. Vintage, Penguin Random House UK, Dublin, Ireland.

Hargittai, E., Marwick, A., 2016. "What can I really do?" Explaining the privacy paradox with online apathy. Int. J. Commun. 10.

Houghton, D.J., Joinson, A.N., 2010. Privacy, social network sites, and social relations. J. Technol. Hum. Serv. 28 (1), 74–94. http://dx.doi.org/10.1080/15228831003770775.

Introna, L.D., 1997. Privacy and the computer: Why we need privacy in the information society. Metaphilosophy 28 (3), 259–275. http://dx.doi.org/10.1111/1467-9973.00055.

ISO Central Secretary, 2016. Systems and Software Engineering — Systems and Software Quality Requirements and Evaluation (SQuaRE) — Measurement of Quality in Use. Standard ISO/IEC 25022:2016, International Organization for Standardization, Geneva, CH, URL https://www.iso.org/standard/35746.html.

Jacob, R., Heinz, A., Décieux, J.P., 2014. Umfrage: Einführung in Die Methoden der Umfrageforschung. Oldenbourg Wissenschaftsverlag.

Janssen, H., Cobbe, J., Norval, C., Singh, J., 2020. Decentralized data processing: personal data stores and the GDPR. Int. Data Priv. Law 10 (4), 356–384. http://dx.doi.org/10.1093/idpl/ipaa016.

Jasanoff, S., 2017. Virtual, visible, and actionable: Data assemblages and the sightlines of justice. Big Data Soc. 4 (2), 1–15. http://dx.doi.org/10.1177/2053951717724477.

Jensen, C., Potts, C., 2004. Privacy policies as decision-making tools: An evaluation of online privacy notices. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '04, Association for Computing Machinery, New York, NY, USA, pp. 471–478. http://dx.doi.org/10.1145/985692.985752.

Jiang, X., Hong, J.I., Landay, J.A., 2002. Approximate information flows: Socially-based modeling of privacy in ubiquitous computing. In: Borriello, G., Holmquist, L.E. (Eds.), UbiComp 2002: Ubiquitous Computing. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 176–193.

Jourard, S.M., 1966. Some psychological aspects of privacy. Law Contemp. Probl. 31, 307.

Karegar, F., Pettersson, J.S., Fischer-Hübner, S., 2020. The dilemma of user engagement in privacy notices: Effects of interaction modes and habituation on user attention. ACM Trans. Priv. Secur. 23 (1), http://dx.doi.org/10.1145/3372296.

Kästner, L., Langer, M., Lazar, V., Schomäcker, A., Speith, T., Sterz, S., 2021. On the relation of trust and explainability: Why to engineer for trustworthiness. In: 2021 IEEE 29th International Requirements Engineering Conference Workshops (REW). pp. 169–175. http://dx.doi.org/10.1109/REW53955.2021.00031.

Keymanesh, M., Elsner, M., Parthasarathy, S., 2021. Privacy policy question answering assistant: A query-guided extractive summarization approach. CoRR abs/2109.14638. arXiv:2109.14638.

Khan, U., Wang, L., Subramanian, J., Near, J.P., Song, D., 2020. PrivFramework: A system for configurable and automated privacy policy compliance. CoRR abs/2012.05291. arXiv:2012.05291. URL https://arxiv.org/abs/2012.05291.

Klitou, D., 2014. Privacy-Invading Technologies and Privacy by Design, first ed. T.M.C. Asser Press, The Hague, The Hague, NL, http://dx.doi.org/10.1007/978-94-6265-026-8.

Klopfer, P.H., Rubenstein, D.I., 1977. The concept privacy and its biological basis. J. Soc. Issues 33 (3), 52–65.

Köhl, M.A., Baum, K., Langer, M., Oster, D., Speith, T., Bohlender, D., 2019. Explainability as a non-functional requirement. In: 2019 IEEE 27th International Requirements Engineering Conference (RE). pp. 363–368. http://dx.doi.org/10.1109/RE.2019.00046.

Kokolakis, S., 2017. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. Comput. Secur. 64, 122–134. http://dx.doi.org/10.1016/j.cose.2015.07.002.

Koskinen, J., Knaapi-Junnila, S., Rantanen, M.M., 2019. What if we had fair, people-centred data economy ecosystems? In: 2019 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI). pp. 329–334. http://dx.doi.org/10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00100.

Krishna, A., 2020. Privacy is a concern: An introduction to the dialogue on privacy. J. Consum. Psychol. 30 (4), 733–735.

Kumaraguru, P., Cranor, L.F., 2005. Privacy Indexes: A Survey of Westin's Studies. Institute for Software Research International.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics.

Langer, M., Oster, D., Speith, T., Hermanns, H., Kästner, L., Schmidt, E., Sesing, A., Baum, K., 2021. What do we want from explainable artificial intelligence (XAI)? – A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research. Artificial Intelligence 296, 103473. http://dx.doi.org/10.1016/j.artint.2021.103473.

McCloskey, H.J., 1980. Privacy and the right to privacy. Philosophy 55 (211), 17–38. http://dx.doi.org/10.1017/S0031819100063725.

McDonald, A.M., Cranor, L.F., 2008. The cost of reading privacy policies 2008 privacy year in review. I/S: J. Law Policy Inf. Soc. 4 (3), 543–568.

McDonald, A.M., Reeder, R.W., Kelley, P.G., Cranor, L.F., 2009. A comparative study of online privacy policies and formats. In: Goldberg, I., Atallah, M.J. (Eds.), Privacy Enhancing Technologies. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 37–55.

McKnight, D.H., 2005. Trust in information technology. In: The Blackwell Encyclopedia of Management. Vol. 7 Management Information Systems. Blackwell Publishing, Malden, MA, USA, pp. 329–331, Ch. 20.

Mehdiyev, N., Houy, C., Gutermuth, O., Mayer, L., Fettke, P., 2021. Explainable artificial intelligence (XAI) supporting public administration processes – On the potential of XAI in tax audit processes. In: Ahlemann, F., Schütte, R., Stieglitz, S. (Eds.), Innovation Through Information Systems. Springer International Publishing, Cham, pp. 413–428.

Miles, M.B., Huberman, A.M., 1994. Qualitative Data Analysis: An Expanded Sourcebook, second ed. SAGE Publications Inc., Thousand Oaks, CA, USA.

Moore, A.D., 2003. Privacy: Its meaning and value. Amer. Philos. Q. 40 (3), 215–227.

Mourey, J.A., Waldman, A.E., 2020. Past the privacy paradox: The importance of privacy changes as a function of control and complexity. J. Assoc. Consum. Res. 5 (2), 162–180. http://dx.doi.org/10.1086/708034.

Nagulendra, S., Vassileva, J., 2016. Providing awareness, explanation and control of personalized filtering in a social networking site. Inf. Syst. Front. 18 (1), 145–158. http://dx.doi.org/10.1007/s10796-015-9577-y.

Newell, P.B., 1995. Perspectives on privacy. J. Environ. Psychol. 15 (2), 87–104. http://dx.doi.org/10.1016/0272-4944(95)90018-7.

Nissim, K., Wood, A., 2018. Is privacy *privacy*? Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. 376 (2128), 20170358. http://dx.doi.org/10.1098/rsta.2017.0358.

Nokhbeh Zaeem, R., Anya, S., Issa, A., Nimergood, J., Rogers, I., Shah, V., Srivastava, A., Barber, K.S., 2020. PrivacyCheck v2: A tool that recaps privacy policies for you. In: Proceedings of the 29th ACM International Conference on Information & Knowledge Management. CIKM '20, Association for Computing Machinery, New York, NY, USA, pp. 3441–3444. http://dx.doi.org/10.1145/3340531.3417469.

Notario, N., Crespo, A., Martin, Y.-S., Del Alamo, J.M., Metayer, D.L., Antignac, T., Kung, A., Kroener, I., Wright, D., 2015. PRIPARE: Integrating privacy best practices into a privacy engineering methodology. In: 2015 IEEE Security and Privacy Workshops. pp. 151–158. http://dx.doi.org/10.1109/SPW.2015.22.

Ohm, P., 2015. Foreword. In: The Architecture of Privacy, first ed. O'Reilly Media, Inc., Sebastopol, CA, USA, pp. 9–11.

Papenmeier, A., Englebienne, G., Seifert, C., 2019. How model accuracy and explanation fidelity influence user trust in AI. In: Proceedings of the IJCAI 2019 Workshop on Explainable Artificial Intelligence (XAI). pp. 94–100.

Parkins, D., 2017. The world's most valuable resource is no longer oil, but data. Econ. 6, URL https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data.

Patil, V., Shyamasundar, R., 2019. Is privacy a myth for facebook users? In: Proceedings of the 16th International Joint Conference on e-Business and Telecommunications - SECRYPT,. INSTICC, SciTePress, pp. 510–516. http://dx.doi.org/10.5220/0008018805100516.

Pentina, I., Zhang, L., Bata, H., Chen, Y., 2016. Exploring privacy paradox in information-sensitive mobile app adoption: A cross-cultural comparison. Comput. Hum. Behav. 65, 409–419. http://dx.doi.org/10.1016/j.chb.2016.09.005.

Petronio, S., 2002. Boundaries of Privacy: Dialectics of Disclosure. State University of New York Press.

Pieters, W., 2011. Explanation and trust: what to tell the user in security and AI? Ethics Inf. Technol. 13 (1), 53–64. http://dx.doi.org/10.1007/s10676-010-9253-3.

Pollach, I., 2007. What's wrong with online privacy policies? Commun. ACM 50 (9), 103–108.

Pötzsch, S., 2009. Privacy awareness: A means to solve the privacy paradox? In: Matyáš, V., Fischer-Hübner, S., Cvrček, D., Švenda, P. (Eds.), The Future of Identity in the Information Society. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 226–236.

Rana, O., Weinman, J., 2015. Data as a currency and cloud-based data lockers. IEEE Cloud Comput. 2 (2), 16–20. http://dx.doi.org/10.1109/MCC.2015.46.

Rantanen, M.M., 2019. Towards ethical guidelines for fair data economy - thematic analysis of values of Europeans. In: Rantanen, M.M., Koskinen, J. (Eds.), Proceedings of the Third Seminar on Technology Ethics 2019. ceur-ws.org, Turku, Finland, pp. 27–38.

Reidenberg, J.R., Breaux, T., Cranor, L.F., French, B., Grannis, A., Graves, J.T., Liu, F., McDonald, A., Norton, T.B., Ramanath, R., 2015. Disagreeable privacy policies: Mismatches between meaning and users' understanding. Berkeley Technol. Law J. 30 (1), 1–88.

Renaud, K., Gálvez-Cruz, D., 2010. Privacy: Aspects, definitions and a multi-faceted privacy preservation approach. In: 2010 Information Security for South Africa. pp. 1–8. http://dx.doi.org/10.1109/ISSA.2010.5588297.

Richardson, A., Rosenfeld, A., 2018. A survey of interpretability and explainability in human-agent systems. In: Proceedings of the IJCAI/ECAI Workshop on Explainable Artificial Intelligence (XAI 2018). pp. 137–143.

Rubenfeld, J., 1989. The right of privacy. Harv. Law Rev. 102 (4), 737–807.

Rudolph, M., Feth, D., Polst, S., 2018. Why users ignore privacy policies – A survey and intention model for explaining user privacy behavior. In: Kurosu, M. (Ed.), Human-Computer Interaction. Theories, Methods, and Human Issues. Springer International Publishing, Cham, pp. 587–598.

Saldaña, J., 2013. The Coding Manual for Qualitative Researchers, second ed. SAGE Publications Inc., Thousand Oaks, CA, USA.

Schafer, B., Edwards, L., 2017. "I spy, with my little sensor": fair data handling practices for robots between privacy, copyright and security. Connect. Sci. 29 (3), 200–209. http://dx.doi.org/10.1080/09540091.2017.1318356.

Schneier, B., 2015. Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World, first ed. W. W. Norton & Company, New York, NY, USA.

Schomakers, E.-M., Lidynia, C., Müllmann, D., Ziefle, M., 2019. Internet users' perceptions of information sensitivity – insights from Germany. Int. J. Inf. Manage. 46, 142–150. http://dx.doi.org/10.1016/j.ijinfomgt.2018.11.018, URL https://www.sciencedirect.com/science/article/pii/S0268401218307692.

Seaman, C., 1999. Qualitative methods in empirical studies of software engineering. IEEE Trans. Softw. Eng. 25 (4), 557–572. http://dx.doi.org/10.1109/32.799955.

Senarath, A., Arachchilage, N.A.G., Slay, J., 2017. Designing privacy for you: A practical approach for user-centric privacy. In: Tryfonas, T. (Ed.), Human Aspects of Information Security, Privacy and Trust. Springer International Publishing, Cham, pp. 739–752.

Sheth, A., Gaur, M., Roy, K., Faldu, K., 2021. Knowledge-intensive language understanding for explainable AI. IEEE Internet Comput. 25 (5), 19–24. http://dx.doi.org/10.1109/MIC.2021.3101919.

Smart, M.A., 2021. Addressing privacy threats from machine learning. In: NeurIPS 2021. URL https://www.tib.eu/de/suchen/id/arxiv3A5a8fe7c981fb6e3f3752b017c19e3a0296909499.

Smith, R.E., 2000. Ben Franklin's Web Site: Privacy and Curiosity from Plymouth Rock to the Internet. Privacy Journal.

Soe, T.H., Nordberg, O.E., Guribye, F., Slavkovik, M., 2020. Circumvention by design - Dark patterns in cookie consent for online news outlets. In: Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society. Association for Computing Machinery, New York, NY, USA, pp. 1–12.

Sudman, S., Bradburn, N.M., 1982. Asking Questions: A Practical Guide to Questionnaire Design. Jossey-Bass Inc., U.S..

Tjoa, E., Guan, C., 2021. A survey on explainable artificial intelligence (XAI): Toward medical XAI. IEEE Trans. Neural Netw. Learn. Syst. 32 (11), 4793–4813. http://dx.doi.org/10.1109/TNNLS.2020.3027314.

Tsai, J., Cranor, L.F., Acquisti, A., Fong, C.M., 2006. What's it to you? A survey of online privacy concerns and risks. NET Inst. Work. Pap. 06 (29), 1–20. http://dx.doi.org/10.2139/ssrn.941708.

Tun-Min, C., King, J., King, N.J., 2016. Privacy versus reward: Do loyalty programs increase consumers' willingness to share personal information with third-party advertisers and data brokers? J. Retail. Consum. Serv. 28, 296–303. http://dx.doi.org/10.1016/j.jretconser.2015.01.005.

Utz, C., Degeling, M., Fahl, S., Schaub, F., Holz, T., 2019. (Un)informed consent: Studying GDPR consent notices in the field. In: Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security. CCS '19, Association for Computing Machinery, New York, NY, USA, pp. 973–990. http://dx.doi.org/10.1145/3319535.3354212.

Velecky, L.C., 1978. The concept of privacy. In: Privacy. John Wiley & Sons, New York, NY, USA, pp. 13–34, Ch. 2.

Warren, S.D., Brandeis, L.D., 1890. The right to privacy. Harv. Law Rev. 4 (5), 193–220.

Westin, A., 2002. Privacy on and off the internet: What consumers want. In: Privacy and American Business. Harris Interactive, New York, NY, USA, pp. 1–126.

Westin, A.F., 2015. Privacy and Freedom. ig Publishing.

Wieringa, J., Kannan, P., Ma, X., Reutterer, T., Risselada, H., Skiera, B., 2021. Data analytics in a privacy-concerned world. J. Bus. Res. 122, 915–925. http://dx.doi.org/10.1016/j.jbusres.2019.05.005.

Wilkowska, W., Offermann-van Heek, J., Colonna, L., Ziefle, M., 2020. Two faces of privacy: Legal and human-centered perspectives of lifelogging applications in home environments. In: Gao, Q., Zhou, J. (Eds.), Human Aspects of IT for the Aged Population. Healthy and Active Aging. Springer International Publishing, Cham, pp. 545–564.

Wirth, J., Maier, C., Laumer, S., Weitzel, T., 2021. Laziness as an explanation for the privacy paradox: a longitudinal empirical investigation, Laziness and the privacy paradox, Internet Research. Internet Res. 32 (1), 24–54. http://dx.doi.org/10.1108/INTR-10-2019-0439, URL https://www.tib.eu/de/suchen/id/emerald3Adoi7E10.1108252FINTR-10-2019-0439.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M.C., Regnell, B., Wesslén, A., 2012. Experimentation in Software Engineering, first ed. Springer, Berlin, Heidelberg, Heidelberg, Germany, http://dx.doi.org/10.1007/978-3-642-29044-2.

Woodruff, A., Pihur, V., Consolvo, S., Brandimarte, L., Acquisti, A., 2014. Would a privacy fundamentalist sell their DNA for $1000...if nothing bad happened as a result? The westin categories, behavioral intentions, and consequences. In: 10th Symposium on Usable Privacy and Security (SOUPS 2014). USENIX Association, Menlo Park, CA, pp. 1–18.

Wu, B., Li, J., Yu, J., Bian, Y., Zhang, H., Chen, C., Hou, C., Fu, G., Chen, L., Xu, T., Rong, Y., Zheng, X., Huang, J., He, R., Wu, B., Sun, G., Cui, P., Zheng, Z., Liu, Z., Zhao, P., 2022. A survey of trustworthy graph learning: Reliability, explainability, and privacy protection. http://dx.doi.org/10.48550/ARXIV.2205.10014.

Yao, M.Z., 2011. Self-protection of online privacy: A behavioral approach. In: Privacy Online: Perspectives on Privacy and Self-Disclosure in the Social Web. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 111–125. http://dx.doi.org/10.1007/978-3-642-21521-6_9.